

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

\*\*\*\*\*

Run on: Sat Apr 10 09:18:04 1999; MasPar time 2266.50 Seconds  
1508.725 Million cell updates/sec

\*\*\*\*\*

Title: >US-08-878-177-1  
Description: (1-1447) from US08878177.seq  
Perfect Score: 1447  
N.A. Sequence: 1 gaatttcgcgcgaagaataa.....gaaagaagcgcgaagaaa 1447  
Comp: ctttaagcgcttgcattat.....cttcttcgcggtcttt

Scoring table: TABLE default  
Gap 6

Mismatch STD : Dbase 0; Query 0

Searched: 602357 segs, 1181590623 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: embd157  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_om 8:em\_ov 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pi  
13:em\_ro 14:em\_un 15:em\_v1

Database: genbank110  
16:gb\_ba1 17:gb\_ba2 18:gb\_htg 19:gb\_in 20:gb\_om 21:gb\_ov  
22:gb\_pat 23:gb\_ph 24:gb\_pi1 25:gb\_pi2 26:gb\_pi1  
27:gb\_pi2 28:gb\_pi3 29:gb\_pi3 30:gb\_pi3 31:gb\_pi3 32:gb\_pi3  
33:gb\_un 34:gb\_v1

Statistics: Mean 11.236; Variance 4.994; scale 2.250

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	949	65.6	1516	21	GGERG	G.gallus ERG mRNA.	0.00e+00
2	799	55.2	3166	26	HUMERG2	Human erg2 gene encodi	0.00e+00
3	653	45.1	3126	26	HUMERG11	Human erg protein (ets	0.00e+00
4	446	30.8	762	26	S72621	EWS.. .erg (transloca	0.00e+00
5	432	29.9	816	29	S66169	Erg-3-immunoglobulin h	0.00e+00
6	382	26.4	1729	29	MMFLI1	Mouse Flt-1 mRNA for r	0.00e+00
7	363	25.1	2916	22	A36461	Sequence 2 from Patent	0.00e+00
8	363	25.1	2938	26	HSHUMFLI	H sapiens HUMFLI-1 mRN	0.00e+00
9	358	24.7	1673	26	S45205	Flt-1-Friend leukemia	4.52e-298
10	355	24.5	4403	21	XLFL16	X.laavis mRNA XLFL1.	3.15e-295
11	353	24.4	13746	27	AP000022	Homo sapiens genomic D	2.47e-293
12	341	23.6	2954	26	HUMERGBFLI	Human ERGB transcripti	5.68e-282
13	337	23.3	588	26	S72865	EWS.. .EWS-ERG-ER	3.46e-278

14	325	22.5	1932	26	HUMFLI1A	Human Flt-1 mRNA, comp	7.63e-267
15	320	22.0	3490	21	CCRNAPFLI	Coturnix coturnix mRNA	4.03e-262
16	318	22.0	3545	21	CCFLIONCO	Coturnix coturnix mRNA	3.11e-260
17	279	19.3	66352	28	HSY17293	Homo sapiens Flt-1 gen	1.68e-123
18	179	12.4	634	19	SULERG	Sea urchin sequence ho	1.85e-120
19	152	10.5	1403	27	HSDNAFEV3	H. sapiens FEV gene, ex	8.54e-106
20	151	10.4	1901	27	HSRNAPFEV	H. sapiens mRNA for FEV	4.50e-103
21	149	10.3	1752	29	RMU91679	Rattus norvegicus ERG	2.35e-100
22	136	10.1	267	26	S73762	EWS.. .erg (reclproc	2.49e-91
23	136	9.4	38006	19	CELR08H4	Caenorhabditis elegans	5.29e-62
24	103	7.1	389	19	DRODERS3A	Drosophila melanogaste	2.97e-60
25	101	6.9	1498	29	MMERG1	M.musculus c-ets-1 gen	2.21e-59
26	100	6.9	328	22	A36462	Sequence 3 from Patent	1.65e-58
27	99	6.8	1475	29	MUSERS1Q	Mouse avian erythrobla	1.65e-58
28	99	6.8	1902	29	MMERS1	M.musculus ets-1 mRNA.	1.65e-58
29	92	6.4	229	26	S72620	EWS.. .Flt1 [human, T	1.94e-52
30	92	6.4	1531	21	CHKCERS21A	Chicken c-ets-1 oncog	1.94e-52
31	91	6.4	2764	21	CHKERS1	Chicken protein p54 (e	1.94e-52
32	91	6.3	211	26	S77574	TL5.. .ERG (transloca	1.41e-51
33	91	6.3	1259	21	XLERS1B	Xenopus laevis c-ets-1	1.41e-51
34	91	6.3	1450	26	HUMERS1A	Human erythroblastosis	1.41e-51
35	91	6.3	1604	26	HSCERS1	Human DNA for c-ets-1	1.41e-51
36	91	6.3	1604	22	196206	Sequence 43 from Paten	1.41e-51
37	91	6.3	1930	21	GGERS2	Chicken mRNA for c-ets	1.41e-51
38	91	6.3	2512	21	XLXFA	X.laavis mRNA XLI-b fo	1.41e-51
39	91	6.3	4991	29	RATERSONCA	Rat proto-oncogene (Et	7.43e-50
40	89	6.2	450	26	HUMERS2	Human Hu-ets-2 gene, h	7.43e-50
41	89	6.2	1678	21	XLFL1B	X.laavis mRNA XLI-b fo	7.43e-50
42	90	6.2	1762	21	GGERS1	Chicken c-ets mRNA for	1.03e-50
43	90	6.2	1856	21	GGERS1	Chicken mRNA for c-ets	1.03e-50
44	89	6.2	3308	19	SUSERS	Strongylocentrotus pur	7.43e-50
45	89	6.2	4032	21	XLERS1A	Xenopus laevis c-ets-1	7.43e-50

ALIGNMENTS

RESULT 1	GGERG	1516 bp	RNA	VRT	27-APR-1995
LOCUS	G.gallus ERG mRNA.				
DEFINITION	X77159				
ACCESSION	g790439				
NID	ERG gene.				
KEYWORDS	chicken.				
SOURCE	Gallus gallus				
ORGANISM	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.				
REFERENCE	1 (bases 1 to 1516) Djordain,P., Dewitte,F., Desbrens,X., Stehelin,D. and Duterque-Cogulland,M. Mesodermal expression of the chicken erg gene associated with precartilaginous condensation and cartilage differentiation Mech. Dev. 50 (1), 17-28 (1995)				
AUTHORS	2 (bases 1 to 1516) Duterque-Cogulland,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-JAN-1994) M. Duterque-Cogulland, CNRS UA 1160, Oncologie Moleculaire, Institut Pasteur, 1 rue Calmette, 59019 Lille, FRANCE				
MEDLINE	95329425				
FEATURES	location/Qualifiers				
source	1. 1516 /organism="Gallus gallus" /db_xref="taxon:9031" /dev_stage="adult" /tissue_type="spleen" /clone_lib="lambda gt10" 63. 1499 /gene="ERG" 63. 1499 /gene="ERG" 63. 1499 /codon_start=1 /db_xref="PDB:g790440"				

/translation: "MASTIKALSVSEFQSLFCATGSPILATENTASSSEYQOT  
SKMSRVPPQDMWLSQPPARVITKMECNQVNGSRNSDDCSVAKGMSVSDNVGM  
NYSYMEKHIIPNNMTINERVIVPADPILMSTDHQWLEMAVKEGLDVLILF  
ONIDKEICKMTKDDFORLTPSYNADILSLHLREPLHLSDVDKALQNSPRL  
MHANRGAATFIPNTSYVPEATORITRPLPEAORSAWTSHTOSKATOPS  
STVKTEDORPLDPOYLOIGTSSRIANPGSGOIOLOMFILELSDSNMCITMEGT  
NGEKMTDPDEVARMRBERKSPNNNDKLSRALRYITDKINIKVHKRRARAFEDRH  
GIALQHPHPRESSMKIPSDLPYMSSTHAPQKNEVAFHPALPVTSSSFFAAPPN  
YMSPTGITYNTRLPAAHMPSHLGTYY"

BASE COUNT 458 a 392 c 327 g 339 t  
ORIGIN

Query Match 65.68; Score 949; DB 21; Length 1516;  
Best Local Similarity 94.78; Pred. No. 0.00e+00;  
Matches 1435; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

Db 1 GAATCCCGCAAGCAATATTATTATAGCAATTATTAGCGATCATATATCTGTATCA 60  
1 gaattcccgcaagcaataattattattagcaattattagcgatcataatacttgatcaca 60  
Oy 61 TTATGCAAGCACTTTAAGAGCAATTATCAGTGTGAGAGAGAGAGAGAGAGAGAGAG 120  
61 ttatgcaagcactttaagagcattatcagtggtgagagagagagagagagagagagagag 120  
Oy 61 ttatgcaagcactttaagagcattatcagtggtgagagagagagagagagagagagagag 120  
Oy 121 AGTGTGCTTACGAGATCGCCCGCCCTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
121 agtgtgcttacgagatcgcccccttgcaagagagagagagagagagagagagagagag 180  
Oy 121 agtgtgcttacgagatcgcccccttgcaagagagagagagagagagagagagagagag 180  
Db 181 AATATGGGCAACATCAAGATGAGCGCGGTTCCCGAGAGAGAGAGAGAGAGAGAGAG 240  
181 aatatgggcaaacatcaagatgagcgcggttcccgagagagagagagagagagagagagag 240  
Oy 181 aatatgggcaaacatcaag 240  
Db 241 CCCCAGCAGAGATTACATTAAGATGAGTGTACCCAAACAGAGTTAAGGTTCAAGGA 300  
241 cccagcagagattacatttaagatgagtgtacccaaacagagtttaagggttcaaggaa 300  
Oy 241 cccagcagagattacatttaagatgagtgtacccaaacagagtttaagggttcaaggaa 300  
Oy 301 ATTCACTGATGACTGCAACCGTGGCAAAAGAGGAAATGGTTAGCAGTTGACACATG 360  
301 attcactgatgactgcacacggtggcaaaagaggaaatggttagcagttgacagcaatg 360  
Oy 301 attcactgatgactgcacacggtggcaaaagaggaaatggttagcagttgacagcaatg 360  
Db 361 TTGGATGAACATGAGGAGCTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
361 ttggatgaacatgagagctacatgagagagagagagagagagagagagagagagagagag 420  
Oy 361 ttggatgaacatgagagctacatgagagagagagagagagagagagagagagagagagag 420  
Oy 361 ttggatgaacatgagagctacatgagagagagagagagagagagagagagagagagagag 420  
Db 421 CCAATGAACGAGAGATTATTGTCCAGCAGATCTACGTTATGAGACAGACATGTAC 480  
421 ccaatgaacgagagattattgtccagcagatctacgttatgagacagacatgttac 480  
Oy 421 ccaatgaacgagagattattgtccagcagatctacgttatgagagagagagagagagag 480  
Oy 481 GGCAGTGCCTGGAGTGGGAGAGTATGTTCTTCAGACGCTGACATCTTGTGT 540  
481 ggcagtgcctggagtgggagagtattgttcttcagacgctgacatcttgtgtgt 540  
Oy 481 ggcagtgcctggagtgggagagtattgttcttcagacgctgacatcttgtgtgt 540  
Db 541 TCCAGACATGATGGGAGAGTGTGTAAATAGCCAAAGATGACTTCCGAGACATCA 600  
541 tccagacatgatgggagagtggtgtaaatagccaaagatgacttccgagacatca 600  
Oy 541 tccagacatgatgggagagtggtgtaaatagccaaagatgacttccgagacatca 600  
Oy 541 tccagacatgatgggagagtggtgtaaatagccaaagatgacttccgagacatca 600  
Db 601 CGCCGAGCTATTAACGCAATATCTCTGTACACCTACACTACCTCAGAGAGATCTC 660  
601 cgccgagcttatTAACGCAATATCTCTGTACACCTACACTACCTCAGAGAGATCTC 660  
Oy 601 cgccgagcttatTAACGCAATATCTCTGTACACCTACACTACCTCAGAGAGATCTC 660  
Oy 601 cgccgagcttatTAACGCAATATCTCTGTACACCTACACTACCTCAGAGAGATCTC 660  
Db 661 TTCCACATTTGACTTCAAGATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720  
661 ttccacatTTGACTTCAAGATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720  
Oy 661 ttccacatTTGACTTCAAGATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720  
Oy 661 ttccacatTTGACTTCAAGATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720  
Db 721 ATGCTAGAAACACAGAGAGCCACTTTTATTTTCCAAATACATCAGTTTACCCAGAG 780  
721 atgctagaaacacagagagccactTTTATTTTCCAAATACATCAGTTTACCCAGAG 780  
Oy 721 atgctagaaacacagagagccactTTTATTTTCCAAATACATCAGTTTACCCAGAG 780  
Oy 721 atgctagaaacacagagagccactTTTATTTTCCAAATACATCAGTTTACCCAGAG 780  
Db 781 CAACGCAAGAAATTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 840  
781 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 840  
Oy 781 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 840  
Oy 781 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 840  
Db 700 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 759  
700 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 759  
Oy 700 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 759  
Oy 700 caacgcaagaaatTAACCAAGGAGATTTTACTTTATGAGAGAGAGAGAGATCAGCCT 759  
Db 841 GGACGAGTCACAGCCATCCACTCAGTCAAAAGCTAACCAACCATCATCTTCAACAGTGC 900  
841 ggacgagtcacagccatccactcagtcAAAAGCTAACCAACCATCATCTTCAACAGTGC 900  
Oy 841 ggacgagtcacagccatccactcagtcAAAAGCTAACCAACCATCATCTTCAACAGTGC 900  
Oy 841 ggacgagtcacagccatccactcagtcAAAAGCTAACCAACCATCATCTTCAACAGTGC 900

Oy 760 ggaagagtcacagccatccactcagtcAAAAGCTAACCAACCATCATCTTCAACAGTGC 819  
Oy 760 ggaagagtcacagccatccactcagtcAAAAGCTAACCAACCATCATCTTCAACAGTGC 819  
Db 901 CCAAAACGAG 960  
901 ccaaaacgag 960  
Oy 820 ccaaaacgag 879  
820 ccaaaacgag 879  
Oy 961 GCCGCTTTCGAAATTCAGGAGAGTGGGAGAGATGAGATGAGATGAGATGAGATGAGATG 1020  
961 gccgctTTTCGAAATTCAGGAGAGTGGGAGAGATGAGATGAGATGAGATGAGATGAGATG 1020  
Oy 880 gccgctttcgaatccag 939  
880 gccgctttcgaatccag 939  
Db 1021 TGTGGAG 1080  
1021 tgtggag 1080  
Oy 940 tgcgag 999  
940 tgcgag 999  
Db 1081 TGACAGACCTTCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1140  
1081 tgacagacCTTCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1140  
Oy 1000 tgcag 1059  
1000 tgcag 1059  
Oy 1141 ACTATGACAAATCAGCGGTCGACTTCGCTACTACTATGACAAAATATTTATGACTAAG 1200  
1141 actatgacAAATCAGCGGTCGACTTCGCTACTACTATGACAAAATATTTATGACTAAG 1200  
Oy 1060 actatgacaaatcagcgctgcacttcgctactactatgacaaaatattatgactaag 1119  
1060 actatgacaaatcagcgctgcacttcgctactactatgacaaaatattatgactaag 1119  
Db 1201 TTGATGTTAAACGCTATGCTTCAAAATTTGATTTTCACAGGAGATGCTAGGCTCCACAGC 1260  
1201 ttgatgTTAAACGCTATGCTTCAAAATTTGATTTTCACAGGAGATGCTAGGCTCCACAGC 1260  
Oy 1120 ttcatgtaaaagctatgctcacaatttgatttccagagatcgctcagcgctccagc 1179  
1120 ttcatgtaaaagctatgctcacaatttgatttccagagatcgctcagcgctccagc 1179  
Db 1261 CTCACCTTCGAAATTCATGCTGATGAAATACCATGAGAGAGAGAGAGAGAGAGAGAGAG 1320  
1261 ctcacCTTCGAAATTCATGCTGATGAAATACCATGAGAGAGAGAGAGAGAGAGAGAGAG 1320  
Oy 1180 ctcaacctcagagatcacaatttgatttccagagatcgctcagcgctccagc 1239  
1180 ctcaacctcagagatcacaatttgatttccagagatcgctcagcgctccagc 1239  
Db 1321 ACCATGACACCCCGAGAGATGAACTTTGATAGTCCCATCCCTGCTTTGCCGTTAA 1380  
1321 accatgacACCCCGAGAGATGAACTTTGATAGTCCCATCCCTGCTTTGCCGTTAA 1380  
Oy 1240 accatgacaccccgagagatgaaacttgatgacttcccatccctcgttgcggttaa 1299  
1240 accatgacaccccgagagatgaaacttgatgacttcccatccctcgttgcggttaa 1299  
Db 1381 CCTCATCAGCTTTTCTGCTGCTTCAATCCATCTGGAATTCACCAATGAGAGAGATCT 1440  
1381 cctcatCAGCTTTTCTGCTGCTTCAATCCATCTGGAATTCACCAATGAGAGAGATCT 1440  
Oy 1300 cctcatcagcttttctgctgcttccatcactactgaaatcccaactgagagcactct 1359  
1300 cctcatcagcttttctgctgcttccatcactactgaaatcccaactgagagcactct 1359  
Db 1441 ACCCCATACAGAGGCTGCGACTGCTCATATGCGTCCATCTGGGAGAGAGAGAGAGAG 1500  
1441 accccatACAGAGGCTGCGACTGCTCATATGCGTCCATCTGGGAGAGAGAGAGAGAGAG 1500  
Oy 1360 accccaatacagagctgcagctcctcatatgcttccatcctcagagcactactaag 1419  
1360 accccaatacagagctgcagctcctcatatgcttccatcctcagagcactactaag 1419  
Db 1501 TGGGGAAGAGAAAGAA 1516  
1501 tggggaaagaaagaa 1516  
Oy 1420 tggggaaagaaagaa 1435  
1420 tggggaaagaaagaa 1435

RESULT 2  
LOCUS HUMERG2 3166 bp ss-RNA PRI 08-NOV-1994  
DEFINITION Human erg2 gene encoding erg2 protein, complete cds.  
ACCESSION M17254  
KEYWORDS erg 2 protein.  
SOURCE Human cell line COLO 320, cDNA to mRNA, clone lambda 12.  
ORGANISM Homo sapiens  
Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 3166)  
AUTHORS Rao, V.N., Papas, T.S. and Reddy, E.S.  
TITLE erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation  
JOURNAL Science 237 (4815), 635-639 (1987)  
MEDLINE 87263429  
FEATURES  
location/Qualifiers  
1..3166  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="21q22.3"  
257..1645  
/gene="ERG"  
257..1645  
/gene="ERG"







gene

Db	146	AGTCAAAAGCTGCTACACCATCTCCCTCCACATGCGCCAAACTGAAGACACGCTCTC	205
Oy	784	agtcanaagctaccacacatcatctctcaacagctgcaccaaaagaaagaccagctactc	843
Db	206	ACTTGGATCTTATCGATTCTTGGACCAACAGTATGCGGCTTGCGAAATCCAGGGCAGT	265
Oy	844	agttgattccctatccagatctcttgacgcagccagccgctcttgcaaaaccagggagtg	903
Db	266	GCCAGATCCAGCTTTGGCAGTTCTCTCTGAGAGCTCGTGTGGAGAGCTCCCACTCAGCT	325
Oy	904	ggcagataagcgtatgagcagttccctactctgaggtcttgctgacaggtcccaactcc	963
Db	326	GCATACCTGGGAAGGCACACGGGAGTTCAAGATGACGATCCGACGAGGTGGCC	385
Oy	964	gcatcacctcggagggcacaaagggggagltcaagatgacagaccctgtagaagtgctc	1023
Db	386	GGCGTGGGAGGAGCGAAGAGCAAAACCAACATGAACTGAGTAAGCTAGAGCTAGCGG	445
Oy	1024	ggcgtctgggagagagaaagaaaccaaaccatacgtactgacaaactcagccggtcac	1083
Oy	1084	tcgcgtaccactcgtatgcacaaatatactatgactaagttcactggttaaacgtatagc	1143
Db	506	AGTTGACCTTCCACGGGATGCGCCAGGCGCTCCAGCCCCCGGAGTCATCTCTGT	565
Oy	1144	aatttgatttccagagaaatcgtcagggccctccagccctccacccctcagaaatcatc	1203
Db	566	ACAAGATCCCTTCAGACGCTCCGTCACATGGGCTCTATACAGGCCACCAAGAAATGA	625
Oy	1204	acaaataaacccatcagacatccctccataatgattctcctacatgacacccccaaga	1263
Db	626	ACTTTGTGGCGGCCACCTCCAGCCCTCCCGCTGACATTTTCAGATTTTTCCTGCC	685
Oy	1264	actttgtagctcccatcccccctgcttggcgttaacctataccagttcttgcgcc	1323
Db	686	CAAACCCATCTGGAATTACCACTAGGGGGGTATATACCCCACTAGAGCTCCCAACA	745
Oy	1324	ctaatccatcactgaaatccacacactggaagcatctacccaataaccagcgtccag	1383
Db	746	GCCATATGCTTCTCAT 762	
Oy	1384	ctcatatgcttcccat 1400	
LOCUS	5	816 bp	21-DEC-1993
DEFINITION	Erg-3-immunoglobulin heavy-chain enhancer-binding Ets protein (alternatively spliced) [mouse, pre-B-cell line 22D6, mRNA partial, 816 nt].		
ACCESSION	S66169		
KEYWORDS	g436276		
SOURCE	Mus sp. pre-B-cell line 22D6.		
ORGANISM	Mus sp.		
REFERENCE	Eukaryotae: mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 816)		
TITLE	Rivera,R.R., Stulver,M.H., Steenbergen,R. and Murte,C.		
JOURNAL	gene expression		
MEDLINE	Mol. Cell. Biol. 13 (11), 7163-7169 (1993)		
REMARK	Genbank staff at the National Library of Medicine created this entry [NCBI g13bbs 138523] from the original journal article.		
FEATURES	This sequence comes from Fig. 2.		
SOURCE	Location/Qualifiers		
	1..816		
	/organism="Mus sp."		
	/db_xref="taxon:10095"		
	1..816		

CDs	partial /note="Immunoglobulin heavy-chain enhancer-binding protein" /gene="Erg-3" 1. 816 /partial /gene="Erg-3" /note="Immunoglobulin heavy-chain enhancer-binding protein: This sequence comes from Fig. 2" /codon_start=1 /db_xref="PID:g43627"
BASE COUNT	201 a 279 c 172 g 164 t
ORIGIN	
Query Match	29.9%; Score 432; DB 29; Length 816;
Best Local Similarity	83.9%; Pred. No. 0.00e+00;
Matches 553; Conservative	0; Mismatches 103; Indels 3; Gaps 2
Db 157	GGTGAGCTTTATTTTCCCAATACCTTCACTATATCCCAAGTACGCAAGAATACCA 216
Qy 657	ggaagccatttatttccaaatacaatcagttacccagaagaacgcaagaataaca 716
Db 217	ACTAGGCAGATTTACCTTATGAGCCCTCCCGAGATCAGCGTGCAGCCGACAGCCAC 276
Qy 717	acaagggcagatttaccttaatgacgaagcgaaggatcagcgatgagatgcaagca - 775
Db 277	CTCACCCCTCAGTCCAAAGCTGCTCAGGCATCTCCCTGTGAGTGCCTCCAAACTGAAGAC 336
Qy 776	-tc-ccactcagtcataaagctaccaccaaccatcattccaaagtcgcccataaagaagac 833
Db 337	CAGGCTCTCAGTTAGATCTTACCAGATCTCTGGGACCGACAGTACCCGCTTGCTAAT 396
Qy 834	cagcgctccctcagttacgatacccttaccgaattcttggaccgcacagcagcgcttgcgaat 893
Db 397	CCAGGTAATGCGAGATCCAGATCTGGAGATGCTGCTGCATGATCCGTGACAGACTCC 456
Qy 894	ccaagggatgagcagatacagctatgacgattcctactggaagcttcgttcggaagatcc 953
Db 457	AACCTCAACTGATCACCCTGGGAAGGCACCAACGGGAGTTCAAGATGACAGACCCGGAC 516
Qy 954	aactccaactgcatcacacctgagggagggacaaaatgaggatgcaagatgacgagacctgat 1013
Db 517	GAGTGCTCGGCGCTGGGGGAGAGAGACGCAAGCCCAATGAACATATGACACTC 576
Qy 1014	gaagtgctcgcgcttggggagaggaagcaaacatacgaactatgacaacatc 1073
Db 577	AGCGCGCGCCCTCCGCTACTACTAGACAAACATCATGACCAAGGTGACCGGGAACGCG 636
Qy 1074	agccgtgacattcgcctactactaagcaaaaatctatgataaagttcagtgtaaaagc 1133
Db 637	TACGCTCAAGATTGTGACTTCACAGGGATTGCCAGGCGCTGCAGCCCACTCTCGAG 696
Qy 1134	tatgcctacaataattgattcttccaggaatcgctcaagccctccagctctcaecctccagaa 1193
Db 697	TGCTCCTGTACAAAGTACCTCCGAGCTGCGCATACATGGGCTCTTATCAGCCACCCC 756
Qy 1194	tcatcatgatacaataaccatcagactcctccctacatagttcctctacatgacacaccc 1253
Db 757	CAGAAGATGAACTTGTGTCTCCCAACCCCTCCGCTCTCCAGTCAACATCTTCAGATT 815
Qy 1254	cagaagatgaacttgtatgctcccatcccccctgccttgcgcgtaacctcatcagctt 1312
LOCUS	6
DEFINITION	Mouse, F1-1 mRNA for retroviral integration site.
ACCESSION	X59421
NID	950974

KEYWORDS c-ets-1 gene; ets gene family; flt-1 gene; integration site.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryotes: mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae: Mus.

AUTHORS 1 (bases 1 to 1729)

TITLE Letwin,K.

JOURNAL Direct Submission

1 (bases 1 to 1729)

REFERENCE Submitted (08-MAY-1991) K. Letwin, Samuel Lunenfeld Res Inst at Mount Sinai Hospital, Div of Mol and Developmental Biol, 600 University Avenue, Toronto Ontario M5G 1X5, CANADA

AUTHORS 2 (bases 1 to 1729)

TITLE Ben-David,Y., Giddens,E.B., Letwin,K. and Bernstein,A.

JOURNAL Erythroleukemia induction by Friend murine leukemia virus: Interleukin activation of a new member of the ets gene family, flt-1, closely linked to c-ets-1

REFERENCE Genes Dev. 5 (6), 908-918 (1991)

MEDLINE 91257578

FEATURES

Location/Qualifiers

1..1729

/organism="Mus musculus"

/strain="balb/c"

/db\_xref="taxon:10090"

/germline

/cell\_type="Friend erythroleukemia"

/cell\_line="CB7"

1..1729

/gene="Flt-1"

/evidence="experimental"

1..1729

/gene="Flt-1"

245..1603

/gene="Flt-1"

/codon\_start=1

/product="retrovirus integration site"

/db\_xref="PID:950975"

/db\_xref="SWISS-PROT:P26323"

/translation="WDGTFKALSVSDOSLFDPSAYAAHLPRADMTASGSPDYQ PKHINLPDQEWINQPVRAWKREYDHMNSRESVDCSYSKMKLYGGEADPMNY NSYDEKNGPPPNPTNERIVPADPTLWTOEHRQMLMAIKELYLMIDISFQ NMDKELCKNKDELRATSAVNTFVLSHLSYRESSLAVNTSHSDOSSRLNVE DPSYDVRGAAMNNMNGSLKSPGLSGSCMGTCBDRBPDPQIIGPSSRLAN GSGQIQLMQFLLELSDSANSCTTBEETNEFKITDDEYARKNGEKSPNNMIDK LSRALRYTDKNIMTKVGRYARKFDEHGIAQLAQPHPTETSMKYPDSISYPSYH AQQQKVFNPSPSPVPTSSFFGAASQYWTSPTAGIYPPNSVPRHNTVPSHLGS Y"

SE COUNT 466 a 484 c 436 g 343 t

IGIN

Query Match 26.4% Score 382; DB 29; Length 1729;

Best Local Similarity 67.9% Pred. No. 0.00e+00;

Matches 894; Conservative 0; Mismatches 404; Indels 18; Gaps 11;

Db 245 ATGGACGGGACATTAAGAGGCTGCTGCTGTGTGAGTGAAGATCAGTCCCTTTTGTAT 304

Qy 63 atggacagacacattaaaggaagcatatcagtggtgaggaagaccagtcctgtttgag 122

Db 305 TCAGCATAGGAGCGGACGACCATCTCCCAAGGAGATATGCTTCGGGGAGTCTT 364

Qy 123 tctgctctaggtctg-c--cccaacttgcaaaagacagaatgtacagccctccctccagt 179

Db 365 GACTACGGGACGCCCAAAATCAACCCCTGCCACCGCAGCAGAGAGTG--ATCA-AC 421

Qy 180 gaattctggcaaacatcaaaagtatgugccgcggttccccaagcagaactggtatcaag 239

Db 422 CAGCAGTGAAGAGTCAATGTCAAGCGGAGTATGACC--AC-ATG--AATGATCCAG 475

Qy 240 ccccccgcagagttacacattaaagtgtggttaaccacaacaggttaagtgttaag 299

Db 476 GAGTTCGGGTGAGCTGTCAAGAAATTAACAAGCTGGGCGGAGCGGAGCC 535

Qy 300 aattcaactgtatgactgcagcgtggcaaaagaggaagaaatggttcaagcattcaagaat 359

Db 536 AACCCATGAATATAAAGTACTGATGATGAAGAAAGCGCCCTCTCCCACTG 595

Qy 360 gtgtgatagaactaagaagctacatgaaagaagcatatccgctccu---atatg 416

Db 596 ACCACCAAGAAAGGAGTGTGTTGCTGTGAGACCCCAACATGTGACACAGGAGAC 655

Qy 417 acaaccaatgaacgaagattatgtgtccagcagaccctcaacgttatgtgcuacagcat 476

Db 656 GTTCGACAGTGGCTGAGTGGCTTTAAAGAAATACGATTGATGAGATTCACACTTCC 715

Qy 477 gtaagcagatgtgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 536

Db 716 TTCTTCAGAACATGTGATGGCAAGGAATTGTGTTAAATGAACAGAGAGACTTCTCCGA 775

Qy 537 tgttccagaacattgagtgaggaagagtggtgttaaatgacaaagagacttccagaga 596

Db 776 GCCACTCCGCTTACACACAGAAATGCTGTTGTGCGACCTCAGTTACTCTACAGGAAGT 835

Qy 597 ctacagccgaagctataacgcaagatatacctcctgtcacacctaacttccagagagaga 656

Db 836 TCA-CTGCTGGCCTATAACACAACTCCCATACAGACCAAGTCTTCAGC-ACGAAATGCA 893

Qy 657 gtagcgaactttatcttccaaatacatcagttacccaagaagcaacgcaagaatuaa 716

Db 894 AGGAAGACCTTCTTATGACTGTCTCAG-GAGAGAGCATGGAACATTAATGAACTCT 952

Qy 717 acaagccagattacattatgacagagagatcagcgtgtgagcagtcacagccat 776

Db 953 GGCCTCAACAAAGTCCTCTCTGTGAGAGTACAGACCACTTGGCAACAAACACTGAGCAG 1012

Qy 777 cccactcagtcacaaagctac 836

Db 1013 CGGCCCCAGCCAGATCTTATCAGATCTCTGAGGCGCCCAACACAGCGCGCTAGCAACCT 1072

Qy 837 cgtctcagtgtaagatccctatcagatcttcttgacagcagcagcagcgtctcguaatca 896

Db 1073 GGGAGTGGGAGATCCAGCTGTGCGAGTTTCTCTGGAATCTACTCCGACAGCGCCAC 1132

Qy 897 gggagtgagcagatcagatcagatcagatcagatcagatcagatcagatcagatcagatc 956

Db 1133 GGCACCTATACACTGGGAGGAGGACCAACGAGGAGTCAAAATGACGAGACCTGATGAG 1192

Qy 957 tccaacttcatacactctggagagggcacaaatggagagttcaagaatgagcagactgaa 1016

Db 1193 GTGGCAGGCGCTGGGAGAGGAGGAGCAAGCCCAACATGAATTAAGACAGCTGAGC 1252

Qy 1017 gtgctcagcgtgtgtggag 1076

Db 1253 CGGCGCTCCGATCTACTATGACAAAATTAATGACCAAGTGATGCAAAAGTAT 1312

Qy 1077 cgtgacacttcgtaactactatgacaanaatattatgactaaagtatcagtgtaaacgtat 1136

Db 1313 GCGTCAAGTTGATTCATGAGGCAATGCCAGGCGCTCGAGCCCAACCAAGAGACA 1372

Qy 1137 gcttacaatttgatttccaaaggaatcgcaagcctccagcctccacccctccagaaatca 1196

Db 1373 TCCATGTCAAGTATCCCTGTGATCTCTCAATGATGCTTCTTCAATGAGCCCATACAG 1432

Qy 1197 tccatgtcaaaataac 1256

Db 1433 AAGGTAACTTTGTCCGCTCTCACCCATCTCCATGCTGTCACCTCTCCAGCTTCTT 1492

Qy 1257 aagatgaacttgtatgctcccaatcccccctgcttgcccgtaacctatcaagcttttt 1316

Db 1493 GGAGGAGATCAATATGAGACCTGCCCACTGTGGAGATCTATCCAAACCCAG 1548

Qy 1317 gctgccttaattccactatgaaatcaccaacttgagagcattccccaataacag 1372

RESULT 7 A36461 2916 bp DNA PAT 05-MAR-1997

LOCUS

DEFINITION Sequence 2 from Patent WO9323549.

Db	770	AAAGAACCCTTCTTATACACACAGCANA -IGAGAGAGCATGGGCAATTAACATGATTTCT	828
Oy	717	ACAAGGCCAAGATCTACCTTACgtaagaaacgagagagatcaagcgttgagcagatcacagccat	776
Db	829	GGCCTCAACAAAGTCCTCCCTTTGGAGGGGACACAAGCATGACATGAATATACAGACAA	888
Oy	777	cccatctagtcataaagctaccccaacatcatctctcaacagtgcccaaaacgaagaccac	836
Db	889	CGGCCACGACAGATCGCTATACAGATCCTGGGCCCGGACGACAGTGGCTATGACCAACCT	948
Oy	837	cgctctcagtagatccttaccagttcttggaccgacagcagcgcgtcttgcaatcca	896
Db	949	GGAAACGGGGAGATCCACCTGTCGGCAATTCCTCCGAGAGTCGTCCTCGACAGGCGCAAC	1008
Oy	897	ggagatgggcagatacagctatgaaagttctctactggagcttcctgtagagcttccaac	956
Db	1009	GCCACCTTATACCTCGAGGAGGGGACCAACGGGGAGTTGCAAAATAGCAGGACCCGATGAG	1068
Oy	957	tccactatgcatcactctggaggggcccacaaatgggagatcacaagatgacagaccctgaa	1016
Db	1069	GTGGCCAGCGCGTGGGGCGAGCGGAAAGCAAGCCCAATGAATTACGACAGCTGAGC	1128
Oy	1017	gtgctcggcgcttgggagagaggaagcaaacctaactatgaaactatgaaacactcagc	1076
Db	1129	CGGGCCCTCGGTATTACTATGATGATTAATAACATTTATGACCAAGTGCAGCGCAAAAGATAT	1188
Oy	1077	cgtagaactcgtcactactatgacaaaatatctatgataaagttcatatgtaaacgcat	1136
Db	1189	GCTTCAAAATTTGACTTCCACAGGGCATGGCCGAGCTGTGCAGCCACATCCGACGAGTCG	1248
Oy	1137	gcttcaaatattgattcttccagagaaatcgctcgaagccctccagcctctaacctccagatca	1196
Db	1249	TCCATGTCACAGTACCTTCTTGACATCTCTTACATGCTTCTTACCATGCCCCACGACAG	1308
Oy	1197	tccatgtaacaaatcccacatcagacctcccctacatagttcctacatgacacacccag	1256
Db	1309	AAGGAGACTTTGTCCTCCCTCCCATCATCCATCCATAGCTTGACCTTCACAGTTCTTT	1368
Oy	1257	aagatgaactttgttagctccccaatcccctgtcttggccgaaactatcagactctttc	1316
Db	1369	GGAGCCGCATCAACATCTGAGACCTCCGCCACGGGGGGAATCTAACCCMACCCCA	1423
Oy	1317	gtcgccctaatccatactagtgaattcaccaacttggagcatcttaccaccaatacca	1371
RESULT	8		
LOCUS	HSNUMFLI	2938 bp	RNA
DEFINITION	H.sapiens HUMFLI-1 mRNA.		
ACCESSION	X67001.544250		
NID	932529		
KEYWORDS	FLI-1 gene homologue.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homalidae; Homo.		
AUTHORS	Delattre,O.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (26-MAY-1992) O. Delattre, lab. de Genet. des Tumeurs. Inst. Curie, 26 rue D'Ulm, 75231 Paris Cedex, FRANCE		
AUTHORS	Delattre,O., Zucman,J., Plouhastel,B., Desmaze,C., Melot,T., Peער,M., Kovar,H., Jobert,I., de Jong,P., Rouleau,G., Aurias,A and Thomas,G.		
TITLE	Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours		
JOURNAL	Nature 359 (6391), 162-165 (1992)		
MEDLINE	92396239		
FEATURES	location/Qualifiers		
SOURCE	1..2938		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		



				/organism="Homo sapiens"	
				/db_xref="taxon:9606"	
gene				134..1492	
				/note="Friend leukemia integration 1"	
				/gene="Fli-1"	
CDS				134..1492	
				/gene="Fli-1"	
				/note="Friend leukemia integration 1, this sequence comes from Fig. 1A"	
				/codon_start=1	
				/db_xref="PID:9257354"	
				/translation="MDCTKEALSYSDSDLPDASGAANHLPKADMTASGPDYQ PHKINLPQOEWINQPRVNVKREIYDHMGSRSPVDCSVSKSLVGGESPPMY NSTWDEKNGPPPMPTTNERVYVADPTLMTQEHVQWLEMAKEYLSMEIDISFQ NMDGKELCKNNKEDFLRATLLYNTVELVLSHSYRESLLAPDPTSDOSSRLSVE DPSYDVRGAMGNMNSGLNKSPLEGAGTISKNTOROPDPTSHLIGTSSRLAMP GSGOIOLOMFLLELSDSANASCTWEGTGEPEFMDPDVARWQGRKSPNNVYK LSRALRYIKNTIKYHKRIYAKFDFHGIQAQLQHPRESSYKIKPSDISYPSIH AHQOKVFPVPHPSMPVSISSFEGAAASQIWTSGGIYPNPNVPHNHPVSHLS YX"	
SE COUNT				428 a 488 c 429 g 328 t	
IGIN					
Query Match				24.7%: Score 358; DB 26; Length 1673;	
Best Local Similarity				68.1%: Pred. No. 4,526-298;	
Matches				898; Conservative 0; Mismatches 396; Indels 24; Gaps 14;	
D	134	ATGACGGGATATTAAAGCGCTGTGCGTGGTGGAGCGACGACCATGCTCTTGAC	193		
O	63	atggcaagcactatataaagaagcattacagctgtagtgaaagaccagctcttgtag	122		
D	194	TCACGCTACGAGCGCGACCCATCTCCCAAGCGGCACATGACTGCTCGGGAGTCT	253		
O	123	tgtgcctcagatcg-c--cccaacctgcaagaagcaaatgacagcctctctccagt	179		
D	254	GACTACGGGACCCCCACAGATCAACCCCTCCACACAGAGAGTGG--ATCA-AT	310		
O	180	gaatatgagcaaacatcaagaatgagccgcgctgcccgagcagagtgtatcacag	239		
D	311	CAGCAGTAGAGGTCAACGTCACAGCGGAGTATGAC---AC-ATG--AATGATCCAG	364		
O	240	ccccggcagaagttaccattatagaatgtagtgaacccaacgaagttaattggtcaga	299		
D	365	GAGTCTCCGCTGAGCTGAGCGTGTAGCAATGACAGCAAGTGTGGCGGAGGACATGC	424		
O	300	aattcactgtatgactgtagcagctggtgcaaaaggaggaataatggttaagatccaacat	359		
D	425	AACCCATGAATCAACAGCTATATGACGAGAGAAATGGCCCCCTCTCCCAACATG	484		
O	360	gttggatgaactatgagactacatgtagaagaagacatctccgccca---atatg	416		
D	485	ACCAACAAGAGAGAGAGTATGCTCCCGCAGACCCCACTGTGACACAGAGCAT	544		
O	417	acaacccaatgagaaagagttatgtgccaagcagctccactacgtttagaagcacacat	476		
D	545	GTGAGGCAATGGCTGGAGTGGCCATAAGAGAGTACACTTGATGATGACACATCC	604		
O	477	gtacggcagtggtctgagctggtgagctgaagagatgtgtctccaagaagtggaactctg	536		
D	605	TTTTTCCAGACATGATGATGAGCAAGCAAGTGTAAATGACAGAGAGATCTCCGC	664		
O	537	ttgttcgaagaacatgtatgagaaagtgtagtaaatgaccaaagaatgacttccagaga	596		
D	665	GCCACACCCCTCTACACACAGAGAGTGTGTGTACACACTTCAGTTACTTACGGA-A	722		
O	597	ctcacgcagagctataaagcagatctctctgtcaccactacactcctaagaagaga	656		
D	723	GTT--CACTGCTGGCTAT--AATACAACCTCCCAACGACCAATCTCAAGATTGAT	778		
O	657	ggagccacttattcttccaataacatcagcttaccagaagcaacgcaagaataaca	716		
D	779	GTCAAAGAGACCTTTTATGACTCACTCAGAGAGAGCTTGGGCAATTAACATGAAT	838		

Oy	717	acaagggccaagttacacttatgacgaacgagggagatcaggtgtggaagatccaagccat	776
D	839	TCGEGCTCAACAANAAGTCTCTCCCTTTGGAGGGCCACAAAGCATCAGTAAGAATACAG	898
Oy	777	ccaactcagt- caaag- -ctaccacaacatcatcttcaacagtgccccaaaacagaac	833
D	899	CAACGGCCCAAGCCAGATCCGTATCAGATCTCTGGGCCGACAGCATGCTGCTAGCCAC	958
Oy	834	caggtcctcagttagatcttaccatcagatctcttgacagcagcagcgcctctgcaat	893
D	959	CCTGGAAGCGGCGATGCCATTCATGGCAATTCCTCCGAGGCTGCTCTCCGACAGGCC	1018
Oy	894	ccagggagtgtggcagatacagctatgtagtccctcagagctctctgtggaagctcc	953
D	1019	AACGCCAGCTGTATCACCTGTGGAGGGGACCAACGGGAGTCAAAATGACGACGCCCAT	1078
Oy	954	aattccaactgtcatcacctcgtgagggcacaatgtggagatccaagaatgacagaccctgt	1013
D	1079	GAGGTGCCAGGCGCTGGGGGCGACGGAAAGCAAGCCCAACATGAATTACGACAGCTG	1138
Oy	1014	gaagtgtcgtcggtgtggtggagagagaaagcaacctaacatgaaactatgacaaactc	1073
D	1139	AGCCGGGCCCTCCGTTATTACTATGATTAACATTTATGACCAAGTGCAGGCCAAAAGA	1198
Oy	1074	agccgtgtcaactctgtctactactatgacaaaataatattagataaagttcatggtataagc	1133
D	1199	TATGCTTACAAATTTGACTTCCACGCGATGTCGCCAGGCTCGACGCCATCCGACCGAG	1258
Oy	1134	tatgcctacaacttgaatttccacaggaatcgtcagagccctccagccctccaccagaa	1193
D	1259	TGCTGCATGTAAGAATACCTTCTGACATCTCTACATGCTTCTTACATGCCACGAG	1318
Oy	1194	tcatcagatgaataataccatcagactccctccatcagatgttcttaccatgacaacccc	1253
D	1319	CAGAAGTGAATTTGTCTCCCTCCCATCATCTTCATGCTGCTGCTCTCCAGCTTC	1378
Oy	1254	cagaagatgaactttgttagtctcccatccctctgttgcgcgttaacctatccagcttt	1313
D	1379	TTTGGAGCGGATCATCATATCTGACCTGCCCAAGCGGGGAATCTACCCACCCCA	1436
Oy	1314	ttgtctgccttaactccatctagatgaattcaccaacttgaggacatctaccaccaatacca	1371
RESULT	10		
LOCUS	XFLIG	4403 bp	RNA
DEFINITION	X.laevius	XFLI.	VRT
ACCESSION	X66979		
NID	5505486		
KEYWORDS	f1l gene; f1l protein.		
SOURCE	African clawed frog.		
ORGANISM	Xenopus laevis		
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;		
	Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;		
	Pipidae; Xenopodinae; Xenopus.		
REFERENCE	1 (bases 1 to 4403)		
AUTHORS	Wolff,C.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-JUN-1992) C.M. Wolff, Institut de Biologie		
	Moléculaire et Cellulaire du C N R S, 15 Rue Rene Descartes, 67084		
	Strasbourg Cedex, FRANCE		
REFERENCE	2 (bases 1 to 4403)		
AUTHORS	Meyer,D., Wolff,C.M., Stiegler,P., Senan,F., Befort,N., Befort,J.J.		
	and Remy,P.		
TITLE	Xl-fil, the Xenopus homologue of the fli-1 gene, is expressed		
	during embryogenesis in a restricted pattern evocative of neural		
	crest cell distribution		
JOURNAL	Mech. Dev. 44 (2-3), 109-121 (1993)		
MEDLINE	94206844		
FEATURES	Location/Qualifiers		
source	1..4403		
	/organism="Xenopus laevis"		
	/db_xref="taxon:8355"		
	/dev_stage="unfertilized eggs"		

```
gene      /clone_11b="cDNA in lambda gt10"
100..1461
/gene="X1f11"
100..1461
/gene="X1f11"
/codon_start=1
/product="Xf11 protein"
/db_xref="pid:9505487"
/db_xref="SWISS-PROT:P41157"
/translation="MDGTRKEALSYVSDQSLFDSAYGASSHLRADMTASANDYQ
PHKINIPPODMINQPMKRVNIKREYHNGSRESVPDCKSINKCKLIGSGENAMTY
TYMDENGPPEPPMTNINERVIVPADPALMSODHYRSLLEIKYGLVEIDCSLFON
IDGKELKMSKEDFLRSISYNTLVLSHLNLRDSSLEAGYNTQAHNDOSRLAKE
DPXYAVRSGKSGMNSSPYTKSPMGSTQNNKGDORSRDPDYQLIGPYSFLAN
PESGQIQLMQFLLELSDSSNASCITWEGNEERKMTDPEVARKWGEKSPKMYND
KLSRLARYDKSIMTKVGRKRYKPFDFHQAOLQHPHPTSNMYKYPSEFSYMSY
HSHQOKNVEFVSHSPMSMPTVSSGFFGATSPLYWNSPANLYPNPNVPRHNPVHQSLG
GFY"
misc_feature 931..1185
/gene="X1f11"
/note="ets-domain"
polya_signal 3449..3455
BASE COUNT 1382 a 836 c 871 g 1314 t
ORIGIN
```

```
Query Match      24.5%; Score 355; DB 21; Length 4403;
Best Local Similarity 68.2%; Pred. No. 3.15e-295;
Matches 749; Conservative 0; Mismatches 340; Indels 9; Gaps 7;
```

```
Db 319 AATGATCCAGGAGTACACCGGTGAGTGTAGCATCAACAAATGTAGCAACTATTGGA 378
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 288 aatggtcaaggaatcaccatgactcagcgtgcgaagaagaggaatagtctagc 347
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 GGGTCGAGGAGGCAATGACTTACACTTACATGATGAAAGAATGGCCCCCTCC 438
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 348 agtctgacaatagtgtgagtaactatgaaactatgaagaagacatattccgcct 407
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 CCAATATATCTACAAATGAAAGAGGTTATTGTACCTGCAGATCCCGTTGTGGAG 498
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 408 ccaaatatgacaacatgaacgaagaagtattgtgcagcagatcctactatgagac 467
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 CAGGACAGTACGAGCCAGTGGGTGAGTGAGCAATTAAGAAATAGTGTGGAAATT 558
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 468 aagaacacatgacgcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 527
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 GACTGCTCCCTCTTCAGATATTTAGTGAAGAACTGTGTAATGAGCAAGAGAT 618
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 528 gcaatctgtgttccagaaatagtgtgaaagagtggtgtataaatagcaaaagatgac 587
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 TTCTCTGGGAGCAGATCCATATATTAACACAGAGTTTACTTTTCATCTCAACTCTC 678
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 588 ttccagaagactcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 647
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 AGGGAATAGTATCATCATTTGGTTATACACTCAGCAGCATACACCAATATATGCGCT 738
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 648 agagaagagggagccacttatttccaaaatacactgattacccagaagaagacgcaa 707
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 TTGACAGCCA-AAGAGATCCCT-CCT-ATGAAGCGGTGAAGAGATCTGGATGGGAAT 795
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 708 agaataaacaacaaagccaaattactatttagcaagcgagagagatcagcggtgagagat 767
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 796 TCCATGAGCTCACTGTCACAAAAAGTCTCCCATGGAGGAACCAAAAATGTGAATAAA 855
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 768 caca-gccatc--ccactcagtcataaagctacccaacc-atacatctcaacagcgccaaa 824
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 856 ACCGGGGATCAGCAAGCTTCTACGCCAGATCCATACCAATATATAGGGCCCAACAAGCAG 915
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 825 aagaagaccagc--gtctcagttagatccttcaatctttagaccgagaccagcagc 881
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 916 ACACGTGCTAACCAAGGAGGTGGCAAAATTCAGCTGTGGCAGTTTCTCTGGAGTGTGCTA 975
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 882 cgtcttgcaaatccagggaggtggagatacagctatgacagttcctacttgagagctctcg 941
```

```
Db 976 TCTGACAGTTCAAACGCTAGTGTATTCATGAGGAGCAGTAAATGAGAGTTCAAAATG 1035
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 942 tcgagacgtcccaactccaactccaactcaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1001
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1036 ACCGACCCAGATGAAGTGAAGCGAGCTGGGTGAAAGGAAAGCAAGCCTTAACATGAAC 1095
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1002 aagaacccgtagaagagtgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1061
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1096 TATGACAAATTAAGCCGAGCAGTAAAGTATTTATGATAAAGTAAATGACTTAAAGTC 1155
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1062 tatgacaaactaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1121
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1156 CATGGTAAACGTTATGCCATACAAATTTGACTTTCACGGCATGTCCAGGACATGAGCCT 1215
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1122 catggttaaacgtatgaccacaaatltgatttccaggaaatgcgcgcgcgcgcgcgcgcgc 1181
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1216 CATCCAACTGATACACTCATGATACAAATATCCTTGTAGTTCTCTTACATGCTCTTAC 1275
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1182 caccctccagaatcattcattgatacaataaccatagacctcccttaacatggtctctac 1241
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1276 CATTCACATCAACAGAGGTTAATTTGTCCCTTCACACCCATCCTCTATGCGGGTAAACA 1335
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1242 catgacaccccccagaagaagtgaactttagcttcccatcccccctgcttgccgtaacc 1301
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1336 TCATCAGGCTTCTTTGGAGCAGCATCTCCTTACTGGAATTCACCAAGTCAAACTTAT 1395
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1302 tcatcagcttcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1361
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1396 CCAAACCCAAATGTGCCA 1413
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1362 cccaataccaggtgcga 1379
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
LOCUS Ap000022 133746 bp DNA PRI 13-MAY-1998
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.2 (Down Syndrome
region), segment 15/15, complete sequence.
ACCESSION Ap000022
NID 93132332
KEYWORDS HTG.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Euthioria;
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 133746)
AUTHORS Hirakawa,M., Yamaguchi,H. and Imai,K.
TITLE Homo sapience 1,296,826bp genomic DNA of 21q22.2 Down Syndrome
region
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 133746)
AUTHORS Hirakawa,M., Yamaguchi,H. and Imai,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1998) to the DDBJ/EMBL/GenBank databases. Miki
Hirakawa, Japan Science and Technology Corporation (JST), Advanced
Databases Department, 5-3, Yonban-cho, Chiyoda-ku, Tokyo, Tokyo
102-0028, Japan (E-mail:mikah@tokyo.jst.go.jp, Tel:81-3-5214-8491,
Fax:81-3-5214-8470)
```

```
FEATURES
source 1..133746
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.2"
```

```
BASE COUNT 38362 a 29023 c 28295 g 38066 t
ORIGIN
```

```
Query Match      24.4%; Score 353; DB 27; Length 133746;
Best Local Similarity 83.7%; Pred. No. 2.47e-293;
Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
```

```
Db 33592 TTATAGTATAGTGGCCAGATGAGGATGATAGCTGTGGGAGGAGCAGTATTTGGGCTA 33651
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 1418 ttatagtagtggtcgaagatgggaagagcatatgacagctgcagcctggttggtgta 1359
```



BASE COUNT	846 a	698 c	683 g	727 t	ORIGIN
Query Match	23.6%; Score 341; DB 26; Length 2954;				
Best Local Similarity	67.9%; Pred. No. 5,68e-282;				
Matches	873; Conservative 0; Mismatches 388; Indels 24; Gaps 14				
Db	173	ATGACGGGACTATTAAAGAGCCTCTGTGCTGTGATGACGACGACCATGCTCTTTTGAC	232		
Qy	63	atggcaagcactataaagaagcattatcagttggtgaggaacagtcctctttttg	122		
Db	233	TCACGGTACGGAGGGGCGACCCCATCTCCCAAGGCCACATGATGCTCTGGGGAGTCT	292		
Qy	123	tgtpcctcagatcg-c-ccacacctgcaaaagacagaataagaaagcttccttcagtt	179		
Db	293	GACTACGGGCGACCCCAACAGATCAACCCCTCCACACACAGCAGAGTGG--ATCA-AT	349		
Qy	180	gaatatggygcaacaatacaataagatgagcccgcggttccccgacgagactggtttacaag	239		
Db	350	CAGCCAGTAGAGGGTCAACGTCMAACGGGAGATGATGACC--AC-ATG--AATGATCCAGG	403		
Qy	240	ccccggcgcaagatgtaaccatttaagatgtagtgaacccaacacagtttaatggtccaag	299		
Db	404	GAGTTCGGGTGAGCTGACGACGCTTAACCAATGCAAGCAAGCTGTGTGGCGGAGGGGAGTCC	463		
Qy	300	aattcaactgtatgactgcagcgctgycacaagaagggaggaataatggtttagcaatcagacaat	359		
Db	464	AACCCATGAGACTCAACACAGCTAATATGAGACAGAAAGATGCGCCCTCCTCCCAACATG	523		
Qy	360	gttggaatgaactatggaagctacatggaagaagaagactatccgcccce--aatatg	416		
Db	524	ACCAACCAACGAGAGAGATCATGCTCCGCCCAACACCCACATCTGTGGACACAGAGCAT	583		

[illegible]



QY 1014 gaagtgctgcgcgttggggaaggaagaaagcaaccctaacatgaaactatgacaaactc 1073  
DB 1178 AGCCGGGCCCCCTCCGTTATTACTATGATTAATAAGATTATGACAAAGTGACAGGCCAAGA 1337  
QY 1074 agcgtgcgaacttcgctactactatgacaaaataatatagtactaaagtcatgtgaaagc 1133  
DB 1238 TATGCTTACAAATTGACTTCCACGAGCATGCCAGCTCTGACGCCACATCCGACGAG 1297  
QY 1134 taagcctacaaatttgatcttcacaggaatcgctcagcctccacagcctccacagaa 1193  
DB 1298 TCGTCACATGACAGTACGCTTGCATCTCCATGACCTGCTTCCAAACATGCCACAG 1357  
QY 1194 tcttcacatgacaaataaccatcagaccctccatcagtgcttcctacacagacccc 1253  
DB 1358 CAGAAGTGAAGTCTTGTCCCTCCCATCATCCATGCTGCTGCTCCAGCTC 1417  
QY 1254 cagaagatgaaactttgtagctcccatcccccctgcttgcgcgtaaacatcagctt 1313  
DB 1418 TTTGGAGCCGCATCACAATCTGGA 1442  
QY 1314 ttgctgcacccatcactactgga 1338

RESULT 13  
LOCUS 572865 588 bp mRNA PRI 24-JAN-1995  
DEFINITION EMS. . EMS-erg-EMS-erg fusion protein type 9e [human, SK-PN-LI cell  
line, mRNA Partial Mutant, 3 genes, 588 nt].  
ACCESSION 572865  
NID 9633777  
KEYWORDS human SK-PN-LI cell line.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 588)  
REFERENCE  
AUTHORS Giovannini, M., Biegel, J.A., Serra, M., Wang, J.Y., Wei, Y.H.,  
Nycum, L., Emanuel, B.S., and Evans, G.A.  
EMS-erg and EMS-Fli1 fusion transcripts in Ewing's sarcoma and  
primitive neuroectodermal tumors with variant translocations  
J. Clin. Invest. 94 (2), 489-496 (1994)  
JOURNAL  
MEDLINE 94314948  
REMARK Genbank staff at the National Library of Medicine created this  
entry [NCBI g15042] from the original journal article.  
This sequence comes from Fig. 2A.  
FEATURES  
source  
1..588  
location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..588  
/partial  
/note="EMS-erg fusion protein type 9e"  
1..588  
gene  
/partial  
/gene="EMS"  
1..588  
/partial  
/gene="erg"  
1..588  
CDS  
/partial  
/gene="EMS-erg"  
/note="EMS-erg fusion protein type 9e; This sequence comes  
from Fig. 2A"  
/codon\_start=1  
/db\_xref="PID:9633778"  
/translation="PTSTPQTSTSYQAPSOYSQSSSTGGQSSGQIQMLLELLS  
DSSNSCITMEGNGEKMTDPDEVARWGERSKENMNYDKLSRLRYUYNIMTK  
VHGKRAYKDFDHGIAQLOPHRPESLSLYKPPDLPMGSMYHAHPQKMNVPVAPHPAL  
PVTSSEFAAPNPNYKNSPTGTYNTRLPSTSHPSH"  
BASE COUNT 145 a 210 c 129 g 104 t  
ORIGIN

Query Match 23.3%, Score 337; DB 26; Length 588;

Best Local Similarity 83.6%; Pred. NO. 3,466-278;  
Matches 419; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
DB 88 AGTGGCAGATCCAGCTTTGGCAGTTCTCTCGAGACTCTCTGAGACAGTCCAACTCC 147  
QY 900 agtggcagataagatagatagagcagttccctacgtgagcttcctggaacagctccaaactc 959  
DB 148 AGCTGCATCAGCTGGAGAGCACCAAGGGAGTTCAAGATGACGAGATCCGACAGGTG 207  
QY 960 aactgcataccctcggagaggaacaaatggagtgtaagaatgacagacccctgtaagtg 1019  
DB 208 GCCCGGCGCTGGGAGAGCGGAAGCAACCAACATGACATGACATGACATGACATGAC 267  
QY 1020 gcttcggtgtgggagagagaaagcaaacatcaatgaaactgaaacactcagcg 1079  
DB 268 GCCCTCGTATCATGATGACAGAACATCATGACCAAGATCCATGAGAGGCTACGCC 327  
QY 1080 gcaactcgctactactatgacaaataatatagtacaaagtccaatgtaagctctgccc 1139  
DB 328 TACAAGTTGACTTCCAGCGGATCGCCAGGCCCTCCAGCCGCCCGGAGTCACT 387  
QY 1140 tacaatttgatttccaggaatcgctcagggccctccagcctcaccctcagaaatcc 1199  
DB 388 CTGTACAAGTACCTCTCAGACCTCCCTGATGAGCTCTCTATGACCCACACGACGAAG 447  
QY 1200 atgtacaaataaccatcagacactccctcactatgattcttcaacatgcaaccccgag 1259  
DB 448 ATGACTTTGTGGCGGCCGCCCTCCAGCCCTCCGCTGATCTCTGATTTTGTGCT 507  
QY 1260 atgaacttgtagctccctccctccctctgcttgccgttaacctcaacagctttttgt 1319  
DB 508 GCCCAACCCATCATCTGATTCACCAAGTGGGGGTATATACCCCAACTAGCTGCC 567  
QY 1320 gccccaatccatcacatcaggaattcaccaactggagatcctacccaataccagcgccga 1379  
DB 568 ACCAGCCATATGCTTCTCAT 588  
QY 1380 gctgcataatgcttcccat 1400

RESULT 14  
LOCUS HUMFL11A 1932 bp mRNA PRI 31-DEC-1994  
DEFINITION Human Fli-1 mRNA, complete cds for two alternate splicings.  
ACCESSION M93255  
NID 9182659  
KEYWORDS Fli-1; alternative splicing; ets oncogene family; transcription  
factor.  
SOURCE Homo sapiens (tissue library: lambda g11 HBL 2T) blood cDNA to  
mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1932)  
REFERENCE  
AUTHORS Hromas, R., May, W., Denny, C., Raskind, W., Moore, J., Maki, R.A.,  
Beck, E., and Klemsz, M.J.  
Human Fli-1 localizes to chromosome 11Q24 and has an aborrrant  
transcript in neuroepithelioma  
JOURNAL  
MEDLINE 93176799  
FEATURES  
source  
1..1932  
location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="HEL"  
/cell\_type="erythro leukemia"  
/tissue\_type="blood"  
/tissue\_lib="lambda g11 HBL 2T"  
/map="11q23-25"  
join(1..401,576..1932)  
/gene="Fli-1"  
/note="alternate"  
1..1932  
/note="alternate"

```
gene
  join(1..401,576..1932)
/gene="FLI-1"
  join(174..401,576..1700)
CDS
  /gene="FLI-1"
  /note="alternate"
  /codon_start=1
  /db_xref="PID:g182660"
  /translation="MDG1KEALSVSDSDSLFSDAYGAANLPRKADMTASGSPDYCG
PHKINLPPOEEMINOPRVKRVYDMMNSESPPVDCSVKSKLVGGGSPNPNNTN
SYDEKNGPPPNMTNERRIVPADATLVTOEHVROMLEMAITEYSLMETIDTSFQNDKELC
MDGKELCKMKNKEDLRATTLTNTYEVLLSHLSYLESSLATNTSHDQSRSLYKED
PSTDSVRGAMNNSGLNKSPPGAGQITISKNEQPDQPTQIIGPTSSRLANPESQIQIWM
SGQIQIWMQFLELLSDSANSCITMEGTNGEFTKMTDEVARARNGORSKRNMYDKLSRLRY
SRLRYTYDKNIMTKVHGKRYAKFDFHGIAQIOLAPHTESSMKYRPSDISYMSYHMOOKVNF
HOKVNFVPHPSMPTVSSSFFGASQYMTSTGCIYFNPVRPHNTHVPSHLGSYX"
  543..1700
/gene="FLI-1"
  /note="alternate"
  /codon_start=1
  /product="FLI-1"
  /db_xref="PID:g182661"
  /translation="MEGGLAGEPRARESPVDCSVKSKLVGGGSPNPNNTN SYDEKN
GPPPNMTNERRIVPADATLVTOEHVROMLEMAITEYSLMETIDTSFQNDKELC
KMKNEDELRATTLTNTYEVLLSHLSYLESSLATNTSHDQSRSLYKEDPSYDVR
RGAMNNSGLNKSPPGAGQITISKNEQPDQPTQIIGPTSSRLANPESQIQIWM
QFLELLSDSANSCITMEGTNGEFTKMTDEVARARNGORSKRNMYDKLSRLRY
YDKNIMTKVHGKRYAKFDFHGIAQIOLAPHTESSMKYRPSDISYMSYHMOOKVNF
VPHPSMPTVSSSFFGASQYMTSTGCIYFNPVRPHNTHVPSHLGSYX"
BASE COUNT      487 a      554 c      519 g      372 t
ORIGIN            chromosome 11q23-25.

Query Match      22.5%   Score 325; DB 26; Length 1932;
Best Local Similarity 68.3%; Pred. No. 7,63e-267;
Matches 719; Conservative 0; Mismatches 322; Indels 12; Gaps 7;

Db 568 GGGCGAGGAGTCTCCGGTGACATCGACGCTTAGCAAAATGACGAAAGCTGTGGCGGAG 627
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 292 ggtcaaggaatcaccatgactgactgacggtgcaaaaggaaggaagaaatgtagtaagct 351
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 628 GCGAGTCCCAACCCCATGAATCAACAGCTATATGAGCAGCAGAGAAGATGGCCCCCTCTCC 687
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 352 cagacaatgcttgatgactaactggaagctacatgtaagaagaagatactccgcctcca- 410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CCAATATGACCCCAACGAGAGAGATGATGTCGCCCAACGCCCAACACTGGTGACAC 747
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 411 --aatatgacaacaaatgaagaagatgctgacgagatcctcactgataatgtagaaga 468
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 748 AGGAGCATGTGAGCAATGAGTGGAGTGGCCATTAAGAGAGATGACGCTTGATGAGATCG 807
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 469 cagacacaatgtaaggaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 528
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 ACACATCTCTTTTCCAGAACATGATGCGCAAGAACTGTGTAAATGAACAAGAGAGACT 867
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 529 acaatctgtgtccagaacatgtagtgaagagtgtagtaaatgaccacaagtgact 588
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 868 TCCTCCGGCGCCACACCCCTCTACACAGGAGAGTGTGTGTACACCTCACTTACTCTTA 927
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 569 tcacagagactcagcgcgagctataacgcgacatcctcctgtccacacccactaccctca 648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 928 GGGG-A-A-GTT--CACTGGCTGGCTAT--ATACAACTCCACACCGACGACATCCTCAG 981
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 649 gagaagagagagagagagagagagagagagagagagagagagagagagagagagagagag 708
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 982 GATTGATGTCTCAAGAGACCTTCTTATGACTCAGTCAAGAGAGAGAGCTTGGGCAATA 1041
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 709 gaataacaacaagcagagttacactatgacagcagcagagagagagagagagagagagag 768
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1042 ACATGATTTCTGGCTCAACAAGTCTCCCTTGGAGGGGACAAAGATCAGTAAATA 1101
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 769 acagacatcccaactagtc-aaaaa--ctaccacaacacatcattcaacagtgccaaaa 825
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1102 ATACAGAGCAAGCGCCCGACGACATCCGTATCAGATCTGTGGCGCGACGACAGTGGCG 1161
```

```
Qy 826 cagaagaccagcgtccctcaatgtagatcccttaccatctcttgagaccgacgacgctc 885
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1162 TAGCCAACTCTGSAAGGGGAGAGATCCAGCTGTGGCAATTCCTCTGGAGTGCCTCCG 1221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 886 ttgcaaatccagggagtgaggcagatacagctcagtgagcttccactgagcttcgtcgg 945
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1222 ACAGGCCCAACGCGACGCTGTATCACTCTGGAGGGGACCAACGAGGAGTTCAAATGACGG 1281
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 946 acagctccaactccaactgcatacctctgaggagggacaaaatgaggagtgtaagatgacag 1005
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1282 ACCCGAGTGAAGTGGCCAGCGCGCTGGGGCAGCGGAAAGCAAGCCCAATGAAATTCG 1341
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1006 accctgataagtgctgctgagtgaggagagagaaagaacaaactaacatgaaactatg 1065
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1342 ACAAGCTAGCGCGGCGCTCCGTTATCTATGATTAATAAACAATTATGACCAAGTGCACG 1401
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1066 acaactcaagcctgacatcctgctactactatgacaaaataatataatgactaaagtcactg 1125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1402 GCAAAAGATATGCTTACAAATTTGACTTCACAGGCAATGCCCAGGCTGTGACGCCACATC 1461
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1126 gtaaacgctatgctcaacaaattgatttccaagagatcgtccagcctccagcctcaacc 1185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1462 CGACGAGTGTCCATGATPACAGTACAGTACCTTGTGACATCTCTACATGCTTCTTACCATG 1521
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1186 ctccgaatcatcatgatacaataaccatacagactccctccataatgagttcctcaatg 1245
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1522 CCCACACAGAGAGTGAACCTTGTCCCTCCCAACATCCCTCCATGCTGCTGATCTCT 1581
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1246 cacaccccccaagaatgaaacttgatgactcccaactccctcgtcttgccgtaacatcat 1305
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1582 CCAGCTCTTGGAGCGCATCATCAATACTGGA 1614
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1306 ccagcttttctgctccctcaatccatctcgtga 1338
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
LOCUS      CCRNAFL1      3490 bp      mRNA      VRT      07-JUL-1998
DEFINITION Coturnix coturnix mRNA for transcription factor FLI, clone 10.1.
ACCESSION Y14773
KEYWORDS   g3269302
SOURCE     fli gene; FLI oncoprotein; transcription factor.
ORGANISM   Coturnix coturnix
            Quail.
            Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
            Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
            Mager,A.M., Grapin-Botton,A., Iadjaali,K., Meyer,D., Wolff,C.M.,
            Stiegler,P., Bonnin,M.A. and Remy,P.
            The avian fli gene is specifically expressed during embryogenesis
            in a subset of neural crest cells giving rise to mesenchyme
            Int. J. Dev. Biol. 42 (4), 561-572 (1998)
JOURNAL    Submitted (09-SEP-1997) P. Stiegler, Centre National de la
MEDLINE    Recherche, Scientifique, U.P.R. 9005 du C.N.R.S., 15 Rue Rene
REFERENCE   Descaerhe, 67084 Strasbourg Cedex, FRANCE
AUTHORS    Stiegler,P.
TITLE      Direct Submission
JOURNAL
COMMENT     Submitted (09-SEP-1997) P. Stiegler, Centre National de la
FEATURES   Related sequence Y14774.
            Location/Qualifiers
            1..3490
            /organism="Coturnix coturnix"
            /db_xref="taxon:9091"
            /chromosome="1"
            /dev_stage="4-day quail embryo"
            /clone.lib="pcDNA1"
            /map="g31-33"
            215..1513
            /gene="f11"
            215..1513
            /gene="f11"
            /gene="f11"
```

```

/function="transcription factor"
/codon_start=1
/product="FLI transcription factor"
/db_xref="PID:e1308265"
/db_xref="PID:g3269303"
/translation="MDGRIKEALSVSDQSLFSDSYGAATHLPKADMTASGNPDYGO
PHKINPLPPODEMTNQPVRYNVRKREYDHNHNSRESPVDCSVKCKSLVAGCTESNPKS
YSTINDERNGPPPNITNDRRYIVPADPLMTQEHVRQMLEWAIKEIGLMEIDTFFQ
NMDGELCKMNRKDFLRKTSLSYLVLSHLSYLRSSSLAYNTPSHTEASSRLATK
EGPVAGTQONVNRKTEQORPOPDYQIIGPTSSRLANPSSQIQLOMFLLELSDSN
ASCITWEGTNGEFTKMDPDEVARRMGRSKRPNMYDKLSRALRYYYDKNIMTKVHGK
RYAYKFDHFHGLAOLQPHPTESSMYKPSDSLMPSYHAHOOKVNFVPHPSMPTVS
SSFGAASPIYTSPSGSTYPRNVRPRNHAHVSPLHSGTY"
BASE COUNT      1026 a      770 c      752 g      942 t
ORIGIN

```

```

Query Match      22.1%; Score 320; DB 21; Length 3490;
Best Local Similarity 79.2%; Pred. No. 4.03e-262;
Matches 434; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

910 AACAGAACCAACGGCCCCAACACGATCCATCAATCCTGGGGCCACACGATGTCG 969
      ||||| || || || ||||| ||||| || || || ||||| || ||
QY  824 aacagagaccagcgcctccagctagatccctatcagatcttggaccgacgaccg 883

DB  970 TCTTGCCATCTCTGGAGTGGGAGATACACGTGGCAGTTCTCTCGAGTTGCTGC 1029
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  884 tcttgaatccagggagtgggagagatacagctatgctacttgagctctgctgc 943

DB  1030 GGACAGTTCCAATGCCAGTGTATCATGGGAAGGACCAATGGGAATTCAAGATGAC 1089
      ||||| ||||| || || ||||| ||||| || || ||||| |||||
QY  944 ggaacagctccaactccaactgcatcaccctgggggacacaaatgggagttcaagatgac 1003

DB  1090 AGACCCAGATGAAGTGGCAAGGGCTGGGAGAGACCAAAAGCAGCCCAACATGATTA 1149
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  1004 agaccctgtagagtggtcgcggtggggagaggaagaaagaaactaaatgaaacta 1063

DB  1150 TGACAAAGCTGAGCCGACCCCTTGATCTACTATGTATGATTAAGACATTATGACAAAGTGCA 1209
      ||||| || ||||| || ||||| ||||| || ||||| ||||| ||
QY  1064 tgacaaactcagccgtgctactgctactatgacaaaatattatgactaaagttca 1123

DB  1210 TGGCAAAAGGTATGCTTACAAATTTGACTTTTCATGGCATTGCCAGGCCCTCCAGCCTCA 1269
      || ||||| ||||| ||||| ||||| || || || ||||| |||||
QY  1124 tggtaaacgtatgcttacaacattgattccaggaatcgctcgaagccctccagcctca 1183

DB  1270 TCCCACTGATGCTCGATGTACAAATCCGTCAGATCTCTCTACATGACCTTCTTACCA 1329
      || ||||| || ||||| || ||||| || ||||| || ||||| ||
QY  1184 ccctccagatcatcatgtaacaaatcacatcagacccctccatcatgattcctaccca 1243

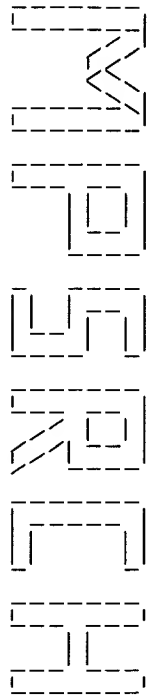
DB  1330 TGGCCACGACGAGAGGTAACCTTGTACCCGACACCCCTTCTTATGCTGTCAATC 1389
      || ||||| ||||| ||||| ||||| || || || ||||| || ||
QY  1244 tgcacacccccagaaatgaaactttagctccccaaccctgcttgcgcgtaaccta 1303

DB  1390 ATCCAGTTCTTTGAGAGCCCTCACCTTATTTGAGCTCCCTCGTGAAGCATTTATCC 1449
      ||||| || ||||| || || ||||| || || ||||| || |||||
QY  1304 atccagcttlttgcgcgcccaatccatcactgaaatccacaactggagggatcctacc 1363

DB  1450 AAACCCCA 1457
      || ||||
QY  1364 caatacca 1371

```

Search completed: Sat Apr 10 10:21:10 1999  
Job time : 3786 secs.



(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MSearch - protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 7 09:46:21 1999; MasPar time 27.16 Seconds  
876,521 Million cell updates/sec

Module output not generated.

Title: >US-08-878-177-4  
Description: (1-478) from US08878177.pep  
Perfect Score: 3467  
Sequence: 1 MASTIKELSVSEDSLE.....IYPNRLPAHMSHLGTY 478

Scoring table: PAM 150  
Gap 11

Searched: 165420 segs, 49795644 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sprenb16  
1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human  
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle  
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified  
13:sp.vertebrate 14:sp\_virus

Statistics: Mean 48.382; Variance 88.436; scale 0.547

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3126	90.2	478 13	Q90837	ERG PROTEIN.	0.00e+00
2	1401	40.4	385 4	Q14319	FLI-1.	9.42e-281
3	1267	36.5	254 4	Q16199	EMS-ERG FUSION PROTEIN	5.03e-250
4	915	26.4	196 4	Q16203	EMS..EMS-ERG-EMS-ER	5.52e-170
5	592	17.1	188 5	Q22355	COSMID T08H4.	4.81e-98
6	555	16.0	208 5	Q18579	SIMILARITY TO ETS DOMA	5.77e-90
7	543	15.7	238 4	Q99581	FEV PROTEIN.	2.35e-87
8	543	15.7	340 11	Q70132	ETS DOMAIN TRANSCRIPT	2.35e-87
9	460	13.3	268 13	Q91744	XEL-A DNA-BINDING PROT	1.84e-68
10	454	13.1	336 6	Q62804	TRANSCRIPTION FACTOR G	3.51e-68
11	449	13.0	539 5	Q26845	ETS HOMOLOGUE.	4.09e-67
12	444	12.8	426 6	Q61805	TRANSCRIPTION FACTOR E	4.75e-66
13	406	11.7	393 5	Q61805	C. ELEGANS LIN-1 (GB:U	5.25e-58
14	406	11.7	441 5	Q27378	LIN-1.	5.25e-58
15	404	11.7	494 13	Q57586	ETS-1 TRANSCRIPT	1.38e-57
16	391	11.3	471 13	Q57566	TRANSCRIPTION FACTOR G	7.41e-55
17	373	10.8	64 4	Q16031	TRANSCRIPTION FACTOR G	4.25e-51
18	336	9.7	76 4	Q16198	EMS..FLI1 (FRAGMENT	1.86e-43
19	326	9.4	663 4	Q99607	MYELOID ELF-1 LIKE FAC	2.06e-41
20	323	9.3	521 4	Q15725	NERF-1A.	8.42e-41

21	323	9.3	533	4	Q15724	NERF-1B.	8.42e-41
22	323	9.3	581	4	Q15723	ETS TRANSCRIPTION FACT	8.42e-41
23	321	9.3	663	4	Q60435	ELF-1 RELATED PROTEIN.	2.15e-40
24	316	9.1	481	5	Q17057	C24A1.2 PROTEIN.	2.32e-39
25	303	8.7	242	5	Q01521	SIMILARITY TP THE DNA-	9.44e-37
26	298	8.6	532	5	Q19695	SIMILAR TO DNA BINDING	9.56e-36
27	296	8.5	81	14	Q64998	(TS1.1 MUTANT) V-ETS (	2.41e-35
28	288	8.3	70	4	Q16324	TLB..ERG (TRANSLOC	9.58e-34
29	279	8.0	300	11	Q70273	EHF (EHF).	5.91e-32
30	269	7.8	314	5	Q01519	SIMILARITY TO THE DNA-	5.58e-30
31	270	7.8	507	5	Q93320	HYPOHETICAL PROTEIN C	3.55e-30
32	265	7.6	348	4	Q99718	EPITHELIAL-SPECIFIC TR	3.41e-29
33	265	7.6	371	4	Q78545	ESB-1B.	3.41e-29
34	265	7.6	331	11	Q35275	ETS TRANSCRIPTION FACT	3.41e-29
35	243	7.0	33	5	Q25573	ETS-1 PROTEIN (FRAGMEN	6.45e-25
36	241	7.0	42	6	Q62767	ERYTHROBLASTOSIS VIRUS	1.56e-24
37	228	6.6	33	5	Q16997	ETS-2 PROTEIN (FRAGMEN	4.74e-22
38	227	6.5	33	5	Q25574	ETS-2 PROTEIN (FRAGMEN	7.32e-22
39	227	6.5	33	5	Q26573	ETS-1 PROTEIN (FRAGMEN	7.32e-22
40	225	6.5	33	5	Q23822	ETS PROTEIN (FRAGMENT)	1.75e-21
41	225	6.5	33	5	Q25066	ETS-1 PROTEIN (FRAGMEN	1.75e-21
42	225	6.5	33	5	Q25127	ETS-1 PROTEIN (FRAGMEN	1.75e-21
43	224	6.5	33	5	Q25159	ETS-1 PROTEIN (FRAGMEN	2.70e-21
44	226	6.5	180	5	Q44138	C50A2.4 PROTEIN.	1.13e-21
45	213	6.1	33	5	Q26070	ETS-1 PROTEIN (FRAGMEN	3.12e-19

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	478 AA.
AC	Q90837			
AD	Q90837			
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)			
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)			
DE	ERG PROTEIN.			
GN	ERG.			
OS	GALLUS GALLUS (CHICKEN).			
OC	ENKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;			
OC	GALLIFORMES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=SPLEEN;			
RC	MEDLINE: 95329425.			
RA	DIORRAIN P., DENITTE F., DESBIENS X., STEHELIN D.,			
RA	DUTERQUE-COQUILLAUD M.;			
RL	MECH. DEV. 50:17-28(1995).			
DR	EMBL: X71159; G790440; -.			
DR	PROSITE: PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE: PS00346; ETS_DOMAIN_2; 1.			
DR	PRAM: PF00178; Ets.			
SO	SEQUENCE 478 AA; 53913 MW; CFE4B9C2 CRC32;			

Query Match 90.2%; Score 3126; DB 13; Length 478;  
Best local Similarity 92.1%; Pred. No. 0.00e+00;  
Matches 441; Conservative 10; Mismatches 26; Indels 2; Gaps 2;

Db	1	MASTIKELSVSEDSLEFCAGSPHLATETMASSSEYGTGTSKSPVPOQDWLSQP	60
Qy	1	MASTIKELSVSEDSLEFCAGSPHLATETMASSSEYGTGTSKSPVPOQDWLSQP	60
Db	61	PARTIMECNPNVNSRNSPDDCSYAKGKMYSSDNGMNGSTMEKHTPPMNTT	120
Qy	61	PARTIMECNPNVNSRNSPDDCSYAKGKMYSSDNGMNGSTMEKHTPPMNTT	120
Db	61	PARTIMECNPNVNSRNSPDDCSYAKGKMYSSDNGMNGSTMEKHTPPMNTT	120
Qy	61	PARTIMECNPNVNSRNSPDDCSYAKGKMYSSDNGMNGSTMEKHTPPMNTT	120
Db	121	NERVIVPADPTLWSTHVMQLEMAVKEYGLPDVDLLFONIDGKELCKMTKDDFORLT	180
Qy	121	NERVIVPADPTLWSTHVMQLEMAVKEYGLPDVDLLFONIDGKELCKMTKDDFORLT	180
Db	121	NERVIVPADPTLWSTHVMQLEMAVKEYGLPDVDLLFONIDGKELCKMTKDDFORLT	180
Qy	121	NERVIVPADPTLWSTHVMQLEMAVKEYGLPDVDLLFONIDGKELCKMTKDDFORLT	180
Db	181	PSYNADILSLHLAYLRTPLPHLTSDDVDKALONSPLMAHARTGATFTFPNTSVYPEA	240
Qy	181	PSYNADILSLHLAYLRTPLPHLTSDDVDKALONSPLMAHARTGATFTFPNTSVYPEA	240
Db	181	PSYNADILSLHLAYLRTPLPHLTSDDVDKALONSPLMAHARTGATFTFPNTSVYPEA	240
Qy	181	PSYNADILSLHLAYLRTPLPHLTSDDVDKALONSPLMAHARTGATFTFPNTSVYPEA	240

Db	241	TORITRPDLEYEOARRSAMTSHHPHQSKATQPSSSVPTTEOQROLOPYOILGPTSS	3000
Oy	241	TORITRPDLEYEOARRSAMTSHHPHQSKATQPSSSVPTTEOQROLOPYOILGPTSS	3000
Db	301	RLANPGSGOIQLOMQLLELLSDSSNSNCITWEGTNGEFKMTDPDEVARRWGERKSKPMNM	3600
Oy	301	RLANPGSGOIQLOMQLLELLSDSSNSNCITWEGTNGEFKMTDPDEVARRWGERKSKPMNM	3600
Db	361	YDKLSRALRITYYDKNIMTKHGKRYAKFDEHGAQALOHPPESSMYKYP-SDLPYMSS	4190
Oy	361	YDKLSRALRITYYDKNIMTKHGKRYAKFDEHGAQALOHPPESSMYKYP-SDLPYMSS	4190
Db	420	YHAPHOKNFEVAPHPALPVTSSEFFAAPNRYMNSPTGCIYPTNRLPAAHMPSHLGYY	4780
Oy	420	PHABQKKNFAPHPALPVTSSEFFAAPNRYMNSPTGCIYPTNRLPAAHMPSHLGYY	4780

RESULT	2	PRELIMINARY;	PRT;	385 AA.
ID	014319			
	014319;			
	01-NOV-1986	(TREMBLREL. 01,	CREATED)	
	01-NOV-1996	(TREMBLREL. 01,	LAST SEQUENCE UPDATE)	
DT	01-JUN-1998	(TREMBLREL. 06,	LAST ANNOTATION UPDATE)	
DE	FLI-1.			
GN	FLI-1.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BLOOD;			
RA	MEDLINE; 93176799.			
RA	HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MARI R.A., BECK E.,			
RA	KLEMSZ M. J.,			
RL	BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).			
DR	EMBL: M93255; G182651; -			
DR	PROSITE; PS00345; ETS_DOMAIN.1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN.2; 1.			
DR	PFAM; PF00178; Ets.			
QO	SEQUENCE 385 AA; 43437 MW; 0820C243 CRC32;			

Query Match	40.48;	Score 1401;	DB 4;	Length 385;
Best Local Similarity	56.48;	Pred. No. 9.42e-281;		
Matches	229;	Conservative	69;	Mismatches 73;
			Indels	35;
			Gaps	21

Db	10	ARSPVDCSVSKSLYVGGGESNPNNYNSYDEMKCPBPNNMTTERRVIVPAADLTITQ	69
Qy	78	SRNSDDCSVAGGAMVSSDDVGNVSNYSIMEKH - IPPNNTTERRVIVPADTLTST	133
Db	70	EHVROMLEMAIKESLIMEIDTSFQNMOMGKELCKNKKEDFLRATLYTEVLLSHLSYLK	122
Qy	137	DHVROMLEMAIKESLYPLPDIILFQNDIGELCKNTKODFOQLTFSYADILSHLHLR	199
Db	130	ESSL - LAYNTTSHT - DOSSRL - SYKE - - D - - - P - - S - Y - DSV - R - - R - G - AMGNM	166
Qy	197	ETPLPHEILSDVDKALQNSPRLMHARNNGATFIEPNTSVYPEAQRITTRPDLPEQAR	251
Db	169	NSG - LNKs - PLYG - GAO - - - TISKATRORPDPDYOLIGPTSSRLANSGSOQLDMQFL	222
Qy	257	RSAMTSHSHPTOSKATLOPSSSTVPTKEQRPOLDYQLIGPTSSRLANSGSOQLDMQFL	311
Db	222	LELSDSANASCITTEGTNGEFGKMTDPDEVARARMGQSKSPMANYDKLSRALRYDYDKNI	288
Qy	317	LELSDSSNSNCITTEGTNGEFGKMTDPDEVARARMGERSKPMANYDKLSRALRYDYDKNI	377
Db	282	MTKYHGKRYAYKEFDHGLAQLAPHPTESSMYKYP - SDISYMSYHAHQKVNYPHPHS	344
Qy	377	MTKYHPRESS - MYKKYPSDLPLYWSSYHGKRYAYKEFDHGLAQLAPHPKOKMFEVAPHP	433
Db	341	SMPLVSSSFEFGAASQWYTS - TGCITPNNVNPVPHNTVNPVSHLSGY	385
Qy	436	ALPVTSSSFEAPNPYWNMSPGGIVPNT - - RLPAHNPVSHLSGY	478

RESULT	3	PRELIMINARY:	PRT:	254 AA.
ID	Q16199			
AC	Q16199			
DT	01-NOV-1996	(TREMBL:REL. 01, CREATED)		
DT	01-NOV-1996	(TREMBL:REL. 01, LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBL:REL. 06, LAST ANNOTATION UPDATE)		
DE	EMS-ERG FUSION PROTEIN TYPE 1E (FRAGMENT).			
GN	EMS-ERG.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUFAROTLA; METAACOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE; 94314948.			
RA	GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYCU M L.,			
RL	EMANUEL B.S., EVANS G.A.;			
RL	J. CLIN. INVEST. 94:489-496(1994).			
DR	EMBL; S72621; G633774; --			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PFAM; PF00178; Ets.			
FT	NON_TER			
SO	SEQUENCE	254 AA;	28577 MR;	596586CD CRC32;

Query Match	36.5%;	Score 1267;	DB 4;	Length 254;
Best Local Similarity	77.7%;	Pred. No. 5.03e-250;		
Matches 178; Conservative			Mismatches 28; Indels 3; Gaps 3;	

Db	27	QONLPEYBRRSAMTGHNPPOSKAAQOPSSVTPKTEDORPOLDPYOIIQGPSRLANP	86
Oy	247	RPDLLEBDARKSAMTSHNHP-QSKAAQOPSSVTPKTEDORPOLDPYOIIQGPSRLANP	305
Db	87	GGSGIOLQMOFLLELTLDSSNSSCITWEGTNEEFKMTDPDEYARRWGERSKSPNNNYKLS	146
Oy	306	GGSGIOLQMOFLLELTLDSSNSNCITWEGTNEEFKMTDPDEYARRWGERSKSPNNNYKLS	365
Db	147	RALRYYYDKNTIMTKVHGKRYAKFDFHGIAOLAQPHPESSLYKYP-SDLPYMGSYAHNP	205
Oy	366	RALRYYYDKNTIMTKVHNPRESS-MYKYPDSLDPYMGSSYGKRAYKFDHFHGIAOLAQPHANP	424
Db	206	QKMFVAPHNPALPYVTSSSFPAANPNYNSTGCIYNTPLRPTSHMPSH	254
Oy	425	QKMFVAPHNPALPYVTSSSFPAANPNYNSTGCIYNTPLRPAAMHPSH	473

```

ID      4
RESULT  4
AC      Q16203
AC      Q16203;
DT      01-NOV-1996 (TREMBLREL, 01, CREATED)
DT      01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)
DT      01-JUN-1998 (TREMBLREL, 06, LAST ANNOTATION UPDATE)
DE      EWS..EWS-ERG-EWS-ERG FUSION PROTEIN TYPE 9E (FRAGMENT).
GN      EWS-ERG.
OS      HOMO SAPIENS (HUMAN) .
OC      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC      EUTHERIA; PRIMATES.
[1]
RN      SEQUENCE FROM N.A.
RP      MEDLINE; 94314948.
RA      GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYC
RA      EMANUEL B.S., EVANS G.A.;
RL      J. CLIN. INVEST. 94;489-496(1994).
DR      EMBL; S72865; G633778; -
DR      PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR      PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR      Pfam; PF00178; Ets.
ET      NON_TER
SQ      SEQUENCE 196 AA; 22268 MW; FBCD632E CRC32;

```

Query Match	26.4%;	Score 915;	DB 4;	Length 196;
Best Local Similarity	73.3%;	Pred. No. 5.52e-170;		
Matches 129;	Conservative	15;	Mismatches 30;	Indels 2;
			Gaps	2

[illegible]

```

DT 01-JUN-1998 (TREMBL:REL. 06, LAST ANNOTATION UPDATE)
DE SIMILARITY TO ETS DOMAINS.
CN C42D8.4.
OC CENOMHABDITIS ELEGANS.
RN EOKAROTA; METAZOA; ACCELLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSOUDH R., ANDERSON K., BAYES C., BERKS M., BONFIELD J.,
RA BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAWTON M.,
RA DEAR S., DU Z., DUBIN R., FAVELLO A., FULTON L., GARDNER A., GREEN P.,
RA HAKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAM J.,
RA KIRSTEY J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C.,
RA MCURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA RIFKEN L., ROOPEA A., SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A.,
RA SONNHAMMER E., STADEN R., SULSTON J., THERRY-MEG G., THOMAS K.,
RA VAUDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., WOHLDMAN P.,
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA HALLSORTH K.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U56966; GI293845; -
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PFAM; PF00178; Ets.
SQ SEQUENCE 208 AA; 22664 MW; 3492097D CRC32;

Query Match 16.0%; Score 555; DB 50; Length 208;
Best Local Similarity 72.6%; Pred. No. 5,77e-90;
Matches 69; Conservative 14; Mismatches 11; Indels 1; Gaps 1

Db 21 GPMALASATGQIQIQMQLLELLADAVNACHAMGSGNCEFLVPDDEVARKGERS 80
QY 236 GPTSSRLANPSSGGIQIQMQLLELLADAVNACHAMGSGNCEFLVPDDEVARKGERS 355
Db 81 KPNMNYDKLSRALRYDDKNTMTKVGQKRYA-YKF 114
QY 356 KPNMNYDKLSRALRYDDKNTMTKVGHPESSMYKX 390

RESULT 7
AC 099581; PRELIMINARY; PRT; 238 AA.
DT 01-MAY-1997 (TREMBL:REL. 03, CREATED)
DT 01-MAY-1997 (TREMBL:REL. 03, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBL:REL. 07, LAST ANNOTATION UPDATE)
DE FEV PROTEIN.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA PETER, COUTURIER J., PACQUEMENT H., MICHON J., THOMAS G.,
RA MAGDELENAI H., DELATRE O.;
RL ONCOGENE 14:1159-1164(1997).
RL EMBL: Y08976; E291921; -
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PFAM; PF00178; Ets.
SQ SEQUENCE 238 AA; 25030 MW; D4057DD0 CRC32;

Query Match 15.7%; Score 563; DB 4; Length 238;
Best Local Similarity 73.7%; Pred. No. 2,35e-87;
Matches 70; Conservative 13; Mismatches 10; Indels 2; Gaps 2

```

[illegible]

```

Oy 257 RSAMTSHHPQSQATOPSSSTVYKTEDQRPQDLPYQILAPTSRLNAPSGQIQLOWFL 316
Db 169 LELTLDSCSGEFSIWGTGDGWEFKLSDPDEVARRRGKKRKNPKMYEKLSGRLYYDKNT 228
Oy 317 LELLSDSSNSNCITWEGNGEFTKTDDEVARRRGKKRKNPKMYEKLSGRLYYDKNT 376
Db 229 IHKT-AGKRYVRFVCDLQSLG 251
Oy 377 MTKVHPPESSMYKYPISDLPIYMSY 400

RESULT 10
AC 062804 PRELIMINARY; PRT; 336 AA.
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE TRANSCRIPTION FACTOR GABP ALPHA SUBUNIT (FRAGMENT).
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA EZASHI T., BALLY A.D., OSTROMSKI M.C., ROBERTS R.M.;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF057717; G3046690; -.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 336 AA; 38033 MW; 8CFE91F CRC32;

Query Match 13.1%; Score 454; DB 6; Length 336;
Best Local Similarity 46.7%; Pred. NO. 3,51e-68;
Matches 63; Conservative 32; Mismatches 34; Indels 6; Gaps 5;

Db 163 PLTAKAIN-SSAKAKAQO-RAPRISGDRSP-GNRTGN--NGQIQLOWFLLELTLDKDA 217
Oy 266 PTQSAKQOPSSSTVYKTEDQRPQDLPYQILAPTSRLNAPSGQIQLOWFLLELSDSSN 325
Db 218 RDCISWVDESEFFKLNPELAAQKQKGNKPNYNEKLSRALRYVYDGMICKVQG-KR 276
Oy 326 SNCITWETNEFMTDPEDEVARRRGKKRKNPKMYEKLSGRLYYDKNTMFKVHPES 385
Db 277 FVYKFCVCDLKTLLIG 291
Oy 386 SMYKYPISDLPIYMSY 400

RESULT 11
AC 026645 PRELIMINARY; PRT; 559 AA.
ID Q26645;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE ETS HOMOLOGUE.
OS STRONGYLOCENTROTUS PURPURATUS (PURPLE SEA URCHIN).
OC EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA;
CC EUDUCHINOIDEA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYONIC;
RA RAO S.K., CHILDS G.;
RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; L19541; G310662; -.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PFM; PF00178; Ets.
SQ SEQUENCE 559 AA; 62052 MW; 0F896532 CRC32;

Query Match 13.0%; Score 449; DB 5; Length 559;
Best Local Similarity 58.0%; Pred. NO. 4.09e-67;
Matches 58; Conservative 18; Mismatches 22; Indels 2; Gaps 2;

```









Db 241 TORITRDPLEEQARRSAMTSHSHPTOSKATOPSSSTVTKEDORPOLDPYQIIGPTSS 300  
QY 214 TORITRDPLEEQARRSAMTSHSHPTOSKATOPSSSTVTKEDORPOLDPYQIIGPTSS 273  
Db 301 RLANGSQIOIOMOFLELLSDSSNSCITWEGTNGEFKMTDPDVARWGERKSKPMNN 360  
QY 274 RLANGSQIOIOMOFLELLSDSSNSCITWEGTNGEFKMTDPDVARWGERKSKPMNN 333  
Db 361 YDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHIGIAOALOPHPPESSMYKYPSDLPYMSY 420  
QY 334 YDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHIGIAOALOPHPPESSMYKYPSDLPYMSY 393  
Db 421 HAHPOKMFVAPHPALPVYSSFFAAPNPYNSPTGCIYPTRLPAHMPSHLGTYY 478  
QY 394 HAHPOKMFVAPHPALPVYSSFFAAPNPYNSPTGCIYPTRLPAHMPSHLGTYY 451

RESULT 2  
ID Q14319 PRELIMINARY: PRT: 385 AA.  
Q14319: 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE FLI-1.  
GN FLI-1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
CC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE: BLOOD;  
RX MEDLINE: 9316799.  
RA HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MAKI R.A., BECK E.,  
KLEMSZ M.J.;  
RL BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PFM: PFO0178; Ets.  
SQ SEQUENCE 385 AA; 43437 MW; 0820C243 CRC32;

Query Match 60.0%; Score 1973; DB 4; Length 385;  
Best Local Similarity 67.8%; Pred. No. 0.00e+00;  
Matches 257; Conservative 68; Mismatches 46; Indels 8; Gaps 6;

Db 10 ARESVDCSVSCSKLVGGESNPNNNSYMDKNGPPNNNTTERRIYVPADATLVTO 69  
QY 78 SRNSDDSCVAKGGMVSSSDVNGMNGSYMEKH-IPPNMTTERRIYVPADATLVTO 136  
Db 70 EHVQMLEMAIKESYIMEIDTSFFQNDGKELCKNNKEDFLRATLLYNTVLLSHLYLR 129  
QY 137 DHVQMLEMAVKEYGLPDVILLFQNDGKELCKMTKDDFORLTPSYNADILLSHLHYLR 196  
Db 130 E-SLLAI-NTSHTDQSSRLSVKEDPSYDSVRGAMGNMNSGLKSPPLGCAQTISN 187  
QY 197 ERGAFFIRPNISYVEAQRITRDPLEEQARRSAMTSHSHPTOSKATOPSSS-TVPKT 255  
Db 188 TEORQPPYQIIGPTSSRLANGSQIOIOMOFLELLSDSANSNCITWEGTNGEFKMTD 247  
QY 256 EDORQPPYQIIGPTSSRLANGSQIOIOMOFLELLSDSANSNCITWEGTNGEFKMTD 315  
Db 248 PDEVARWGERKSKPMNNYDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHIGIAOALOPHP 307  
QY 316 PDEVARWGERKSKPMNNYDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHIGIAOALOPHP 375  
Db 308 TESSMYKPSDLPYMSYHAHQKVFVPPHSSMPVYSSFFGAASQWTIS-TGCIYFN 366  
QY 376 PESSMYKPSDLPYMSYHAHQKVFVPPHSSMPVYSSFFGAASQWTIS-TGCIYFN 435  
Db 367 PNVPRHPTHVPSHLGSYY 385  
QY 436 T---RLPAHMPSHLGTYY 451

RESULT 3  
ID Q16199 PRELIMINARY: PRT: 254 AA.  
Q16199: 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE EMS-ERG FUSION PROTEIN TYPE 1E (FRAGMENT).  
GN EMS-ERG.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
CC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE: 94314948.  
RA GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYCU L.,  
EMANUEL B.S., EVANS G.A.;  
RL J. CLIN. INVEST. 94:489-496(1994).  
DR EMBL: S72621; G633774; -;  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PFM: PFO0178; Ets.  
FT NON-TER 1  
SQ SEQUENCE 254 AA; 28577 MW; 596580CD CRC32;

Query Match 48.7%; Score 1600; DB 4; Length 254;  
Best Local Similarity 93.4%; Pred. No. 0.00e+00;  
Matches 213; Conservative 12; Mismatches 2; Indels 1; Gaps 1;  
Db 27 QONLPYPPRRSAMTGHHPPOSKAOPSPSTYPKTDORPOLDPYQIIGPTSSRLANP 86  
QY 220 RPDLEEQARRSAMTSHSHPT-OSKATOPSSSTVTKEDORPOLDPYQIIGPTSSRLANP 278  
Db 87 GSGQIOMOFLELLSDSSNSCITWEGTNGEFKMTDPDVARWGERKSKPMNNYDKLS 146  
QY 279 GSGQIOMOFLELLSDSSNSCITWEGTNGEFKMTDPDVARWGERKSKPMNNYDKLS 338  
Db 147 RALRYYYDKNIMTKVHGKRAYAKFDFHIGIAOALOPHPPESSLYKYPSDLPYMSYHAHQ 206  
QY 339 RALRYYYDKNIMTKVHGKRAYAKFDFHIGIAOALOPHPPESSLYKYPSDLPYMSYHAHQ 398  
Db 207 KMFVAPHPALPVYSSFFAAPNPYNSPTGCIYPTRLPAHMPSH 254  
QY 399 KMFVAPHPALPVYSSFFAAPNPYNSPTGCIYPTRLPAHMPSH 446

RESULT 4  
ID Q16203 PRELIMINARY: PRT: 196 AA.  
Q16203: 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE EMS-ERG FUSION PROTEIN TYPE 9E (FRAGMENT).  
GN EMS-ERG.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
CC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE: 94314948.  
RA GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYCU L.,  
EMANUEL B.S., EVANS G.A.;  
RL J. CLIN. INVEST. 94:489-496(1994).  
DR EMBL: S72621; G633778; -;  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PFM: PFO0178; Ets.  
FT NON-TER 1  
SQ SEQUENCE 196 AA; 22268 MW; F8CD632E CRC32;

Query Match 38.0%; Score 1248; DB 4; Length 196;  
Best Local Similarity 93.7%; Pred. No. 1.17e-263;  
Matches 164; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

ID	AC	Q22355	PRELIMINARY:	PRF:	188 AA.
Db	22	SSSYGQSSGGIOLIQMOLFLELLSSSSNSCITWGTGEEFMTPDDVARRWGKSKPN	81		
Qy	272	SSRLANPSSGGIOLQMFLELLSSSSNSCITWGTGEEFMTPDDVARRWGKSKPN	331		
Db	82	MNYDKLSRALRYYYDKNIMTKVHKRKYAKKDFFGIAQALOPHPPESSLYKYPEDLPYG	141		
Qy	332	MNYDKLSALALYYYYDKNIMTKVHKRKYAKKDFFGIAQALOPHPPESSMYKYPEDLPYMS	391		
Db	142	SYHAHPQKAMNVAAHPHPPALPYTSSSFFAAPNPYNSPTGGIYPTNRLPTSMPSH	196		
Qy	392	SYHAHPQKAMNVAAHPHPPALPYTSSSFFAAPNPYNSPTGGIYPTNRLPAAMPSPH	446		
RESULT	5	PRELIMINARY:	PRF:	188 AA.	
ID	Q22355	AC	Q22355		
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)			
DT	01-NOV-1996	(TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-JUN-1998	(TREMBLREL. 06, LAST ANNOTATION UPDATE)			
FE	COSMID T08H4.				
FE	T08H4.3.				
OC	CARONSHADITIS ELEGANS.				
OC	EFUAROTA: METAQOA: ACOELOMATES; NEMATODA; SECCERNMENTEA; RHABDITIDA.				
EN	[1]				
EN	SEQUENCE FROM N.A.				
RP	STRAIN-BRISTOL N2;				
RC	MEDLINE: 94150718.				
EX	WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAFTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULLON L., GARDNER A., GREEN P., HARKINS T., HILLIER L., JTER M., JOHNSTON L., JONES M., KERSHAM J., KIRSTEY J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C., MCURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C., RITKEN L., ROOPRA R., SAUNDERS D., SHOWNKEN R., SMALDON N., SMITH A., SONNHAMMER E., STAEDEN K., SUSTON J., THIERRY-MIG J., THOMAS K., VANDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;				
RA	NATURE 368:32-38(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RA	STELLERS L.;				
RL	SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.				
RC	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RA	WATERSTON R.;				
RA	SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.				
EMBL:	U39470: G1041871.				
RC	PROSITE: PS00345; ETS_DOMAIN_1: 1.				
RC	PROSITE: PS00346; ETS_DOMAIN_2: 1.				
DR	PFAM: PF00178; ETS.				
SO	SEQUENCE 188 AA; 21345 MW; C98FD1EC CRC32;				
Db	Query Match	22.2%;	Score 730;	DB 5;	Length 188;
Db	Best Local Similarity	82.8%;	Pred. No. 3,82e-138;		
Qy	Matches	96;	Conservative	8;	Mismatches 11; Indels 1; Gaps 1
Db	5	DPOYIGTSTKSLAHSGSGGTOLQMFLELLSDKRYSEVITWEGTQGEFKLVNDEYAR	64		
Qy	263	DPOYIGTSTKSLANPSSGGIOLQMFLELLSDSSNSCITWGTGEEFMTPDDVARR	322		
Db	65	WGERSKRNMMYDKMSRALRYYYDKNIMAKYHGRKYAKKDFFGIAQALOP-PRAS	119		
Qy	323	WGERSKRNMMYDKMSRALRYYYDKNIMTKVHKRYAKKDFFGIAQALOPHPES	378		
RESULT	6	PRELIMINARY:	PRF:	238 AA.	
ID	Q99581	AC	Q99581		
DT	01-MAY-1997	(TREMBLREL. 03, CREATED)			
DT	01-MAY-1997	(TREMBLREL. 03, LAST SEQUENCE UPDATE)			

Query Match	20.3%	Score 668;	DB 11;	Length 340;
Best Local Similarity	77.3%	Pred. No. 2,08e-123;		
Matches	85;	Conservative	15;	Mismatches 9; Indels 1; Gaps 1
Db	34	GPLSPAYOK-GSGGIQIMQFLLELLADBRANAGCIAMEGSGEFLTPDDEVARRMGERKS	92	
Qy	269	GPTSSRLANPESGOIQIMQFLLELLSDSSNCSNCTWGTGNEFMTDPDEVARRMGERKS	328	
Db	93	KPNMNYDKLSRALRYDYDKNIMSKVHGKRYAYRRDPFGIQAACPPRAHA	142	
Qy	329	KPNMNYDKLSRALRYDYDKNIMTYVHGKRYAYKDFGIAQALPDPHPES	378	
RESULT	7	PRELIMINARY;	PRT;	340 AA.
ID	070132			
AC	070132			
DT	01-AUG-1998	(TREMBLREL. 07, CREATED)		
DI	01-AUG-1998	(TREMBLREL. 07, LAST SEQUENCE UPDATE)		
DI	01-AUG-1998	(TREMBLREL. 07, LAST ANNOTATION UPDATE)		
DE	ETS DOMAIN TRANSCRIPTION FACTOR PET-1.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUDAROTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-PHEOCHROMOCYTOMA;			
RC	MEDLINE; 98127904.			
RA	FYODOROV D., NELSON T., DENEBIS E.;			
RL	J. NEUROBIOL. 34:151-163(1998).			
DR	EMBL; U91679; G3033419; -			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
SO	SEQUENCE 340 AA; 35439 MW; B3214B28 CRC32;			
Query Match	20.3% <td>Score 668;<td>DB 11;<td>Length 340;</td></td></td>	Score 668; <td>DB 11;<td>Length 340;</td></td>	DB 11; <td>Length 340;</td>	Length 340;
Best Local Similarity	77.3% <td>Pred. No. 2,08e-123;</td> <td></td> <td></td>	Pred. No. 2,08e-123;		
Matches	85; <td>Conservative</td> <td>15;<td>Mismatches 9; Indels 1; Gaps 1</td></td>	Conservative	15; <td>Mismatches 9; Indels 1; Gaps 1</td>	Mismatches 9; Indels 1; Gaps 1
Db	137	GPLSPAYOK-GSGGIQIMQFLLELLADBRANAGCIAMEGSGEFLTPDDEVARRMGERKS	195	
Qy	269	GPTSSRLANPESGOIQIMQFLLELLSDSSNCSNCTWGTGNEFMTDPDEVARRMGERKS	328	
Db	136	KPNMNYDKLSRALRYDYDKNIMSKVHGKRYAYRRDPFGIQAACPPRAHA	245	
Qy	329	KPNMNYDKLSRALRYDYDKNIMTYVHGKRYAYKDFGIAQALPDPHPES	378	
RESULT	8	PRELIMINARY;	PRT;	208 AA.
ID	Q18579			
AC	Q18579			
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)		
DI	01-NOV-1996	(TREMBLREL. 01, LAST SEQUENCE UPDATE)		
DI	01-JUN-1998	(TREMBLREL. 06, LAST ANNOTATION UPDATE)		
DE	SIMILARITY TO ETS DOMAINS.			
CN	CA42D.4			
OS	CAENORHABDITIS ELEGANS.			
OC	EUDAROTA; METAFOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.			

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA DEAR S., DU Z., DURBIN R., FAVELL A., FULTON L., GARDNER A., GREEN P.,
RA HAKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAM J.,
RA KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C.,
RA MCURRAY A., MORTMORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA RIEKEN L., ROOPRA A., SANDERS D., SHOWNKEEN R., SMALDON N., SMITH A.,
RA SONNHAMMER E., STADEN R., SULSTON J., THIERRE-MIEG J., THOMAS K.,
RA VAUDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., MOHLDMAN P.,
RA NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA HALLSWORTH K.;
RA SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA WATERSTON R.;
RA SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA PROSITE: PS00345; ETS_DOMAIN_1; 1.
RA PROSITE: PS00346; ETS_DOMAIN_2; 1.
RA PFAM: PF00178; Ets.
RN [5]
RP SEQUENCE 208 AA; 22664 MW; 3492097D CRC32;

Query Match 20.28; Score 664; DB 5; Length 208;
Best Local Similarity 79.48; Pctd. No. 1.84e-122;
Matches 81; Conservative 14; Mismatches 7; Indels 0; Gaps 0;

Db 21 GPMALASATGTGQIQIOMQFLELLADAVANACIAGSGEFLVDPDEVARWGERKS 80
QY 269 GPTSSSLAMPSCQIQIOMQFLELLSDSSNCKITWEGTNGEFGKMDPDEVARWGERKS 328
QY 329 KPMNMYDKLSRALRYYYDKNIMTKVQGRYAYKDFQGLAQ 122
QY 329 KPMNMYDKLSRALRYYYDKNIMTKVQGRYAYKDFQGLAQ 370

RESULT 9
ID 062804 PRELIMINARY: PRT: 336 AA.
AC 062804;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION FACTOR GABP ALPHA SUBUNIT (FRAGMENT).
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA EZASHI T., EALY A.D., OSTROMSKI M.C., ROBERTS R.M.;
RA SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF057717; G3046690; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
RN [2]
RP NON_TER 1
SQ SEQUENCE 336 AA; 38033 MW; 8CFEE91F CRC32;

Query Match 19.88; Score 650; DB 6; Length 336;
Best Local Similarity 41.68; Pctd. No. 3.78e-119;
Matches 101; Conservative 61; Mismatches 61; Indels 20; Gaps 15;

Db 58 EDERLGIPIYDPIQWSTDQVLAHVWVWVWKEFSMTDILHTL-NISGRELCSLSEDEFORV 116
QY 121 NRRRVIVPADPLMTSTDVHROMLEMAVKEVGLPDVILLFQONIDGKELCKMTKDDF-QRL 179
QY 117 -PR--GEILMSHLELR-K---YVLASQ--QOMNE-IYTI-DQP-VQIIPAS-VQSAP 163
```

```

QY 180 TPEYNADILLSHLHYLRREGATFIIPNTSVYPEATQRIITRPDLPEQARRSAMTSHSP 239
QY 164 TAIKAIN-SKAKAYO-RAPRISGEDRSSP-GNRTGN--NGQIQIOMQFLELLTDKAR 218
QY 240 TOSKATOPSSSTVPKTEDQRPQIDPQIIGPTSSRLANGSQIQIOMQFLELLSDSSNS 299
Db 219 DCISWGDGEFRLNPELYAQMGQRKKPTMYNEKLSRALRYYYDGMICKVQGRV 278
QY 300 NCTIWTGTEGFEKMDPDEVARWGERKSPPNNYDKLSRALRYYYDKNIMTKVHGKRYA 359
Db 279 YKF 281
QY 360 YKF 362

RESULT 10
ID 091744 PRELIMINARY: PRT: 268 AA.
AC 091744;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE XEL-A DNA-BINDING PROTEIN (FRAGMENT).
GN XEL-A.
OS XENOPUS LAEVIS (AFRICAN CLAMMED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TOTAL OVARY FROM SINGLE FROG;
RA SALVATI F., MORABITO, MERENDINO, CARNEVALI;
RA SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65167; G517228; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PFAM: PF00178; Ets.
RN [2]
RP DNA-BINDING. 1
FT NON_TER 1
SQ SEQUENCE 268 AA; 30805 MW; 39A1C89D CRC32;

Query Match 15.08; Score 494; DB 13; Length 268;
Best Local Similarity 49.68; Pctd. No. 1.14e-82;
Matches 66; Conservative 28; Mismatches 38; Indels 1; Gaps 1;

Db 110 HDSFSEDYPP-AMPSHKSKGTFKDYVRDAELNMDKPIVPAALAGYTGSPQIOLMOFL 168
QY 230 RSAMTSHSHPTQSKATQPPSSVPTKEDQRPQIDPQIIGPTSSRLANGSQIQIOMQFL 289
Db 169 LELLTQKSCQSFISWTDGDEFEKLSPPDEVARWGERKSPPKMYNEKLSRGLRYYYDKNI 228
QY 290 LELLSDSSNSNCTIWTGTEGFEKMDPDEVARWGERKSPPNNYDKLSRALRYYYDKNI 349
Db 229 IHKTAGRYVYRF 241
QY 350 MKTVHGKRYAYKF 362

RESULT 11
ID 062803 PRELIMINARY: PRT: 426 AA.
AC 062803;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION FACTOR ETS-2 (FRAGMENT).
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA EZASHI T., EALY A.D., OSTROMSKI M.C., ROBERTS R.M.;
RA SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF057716; G3046688; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
```







\*\*\*\*\*  
WISRELE (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

\*\*\*\*\*  
Run on: Sat Apr 10 11:51:41 1999; Maspar time 2419.79 Seconds  
1492.252 Million cell updates/sec  
\*\*\*\*\*

Title: >US-08-878-177-3  
Description: (1-1528) from US08878177.seq  
Perfect Score: 1528  
N.A. Sequence: 1 gaattccgcgcagcaataat.....gaagaagagcccaagaana 1528  
Comp: ctttagcgctctgtattat.....cttccttcgcggtctctt

Scoring table: TABLE default  
Gap 6  
Mismatch SMD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb157  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_om 8:em\_ov 9:em\_ph 10:em\_ph 11:em\_ph 12:em\_ph  
13:em\_ro 14:em\_un 15:em\_vl  
Database: genbank110  
16:gb\_ba1 17:gb\_ba2 18:gb\_htg 19:gb\_in 20:gb\_om 21:gb\_ov  
22:gb\_ph 23:gb\_ph 24:gb\_ph 25:gb\_ph 26:gb\_ph  
27:gb\_ph 28:gb\_ph 29:gb\_ph 30:gb\_ph 31:gb\_ph 32:gb\_ph  
33:gb\_un 34:gb\_vl

Statistics: Mean 11.283; Variance 4.994; scale 2.259

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1516	99.2	1516	21	GGERG	G.gallus ERG mRNA.	0.00e+00
2	552	36.1	816	29	S66169	Erg-3-immunoglobulin h	0.00e+00
3	549	35.9	3166	26	HUMERG2	Human erg2 gene encod	0.00e+00
4	463	30.3	3126	26	HUMERG11	Human erg protein (ets	0.00e+00
5	446	29.2	762	26	S72621	EMS. . .erg (transloc	0.00e+00
6	353	23.1	133746	27	AP000022	Homo sapiens genomic D	2.00e-294
7	337	22.1	588	26	S72865	EMS. . .ERG-ERG-EMS-er	4.33e-279
8	320	20.9	3490	21	CCRNFLI	Coturnix coturnix mRNA	4.49e-263
9	318	20.8	3545	21	CCRNFLI	Coturnix coturnix mRNA	3.55e-261
10	316	20.7	1729	29	MFLI1	Mouse Flt-1 mRNA for r	2.78e-259
11	312	20.4	2916	22	A36461	Sequence 2 from Patent	1.71e-255
12	312	20.4	2938	26	HSUMFLI	H.sapiens HUMFLI-1 mRN	1.71e-255
13	310	20.3	1673	26	S45205	Flt-1-Friend leukemia	1.34e-253

14	295	19.3	1932	26	HOMFLI1A	Human FLI-1 mRNA, comp	2.08e-239
15	293	19.2	2954	26	HUMERGBFLI	Human ERGB transcript	1.62e-237
16	288	18.8	4403	21	XLFLIG	X.laewis mRNA XLFLI	8.56e-233
17	279	18.3	66352	28	HSY17293	Homo sapiens FLI-1 gen	2.67e-224
18	179	11.7	634	19	SUEERG	Sea urchin sequence ho	6.87e-131
19	152	9.9	1403	27	HSDNAFEV3	H.sapiens FEV gene, ex	3.94e-106
20	151	9.9	1901	27	HSRNAMEV3	H.sapiens mRNA for FEV	3.22e-105
21	149	9.8	1752	29	RNU91679	Rattus norvegicus ETS	2.13e-103
22	146	9.6	267	26	S73762	EMS. . .erg (reclproc	1.14e-100
23	136	8.9	162	26	S62204S2	eryp55 (alternatively	1.31e-91
24	136	8.9	162	27	S62206S2	Homo sapiens p55 (erg)	1.31e-91
25	136	8.9	38006	19	CELT08H4	Caenorhabditis elegans	1.31e-91
26	103	6.7	389	19	DRODETS3A	Drosophila melanogaste	3.57e-62
27	101	6.6	1498	29	MMCE1S1	M.musculus c-ets-1 gen	2.03e-60
28	100	6.5	328	22	A36462	Sequence 3 from Patent	1.53e-59
29	99	6.5	1475	29	MUETS1Q	Mouse avian erythrobla	1.14e-58
30	99	6.5	1902	29	MMETS1	M.musculus ets-1 mRNA	1.14e-58
31	91	6.0	211	26	S77574	TLN. . .ERG (transloca	1.04e-51
32	92	6.0	229	26	S72620	XEN. . .FLI1 [human, T	1.41e-52
33	91	6.0	1259	21	XLCE1S1B	Xenopus laevis c-ets-1	1.04e-51
34	91	6.0	1450	26	HUMETS1A	Human erythroblastosis	1.04e-51
35	92	6.0	1531	21	CHKERS21A	Chicken c-ets-1 oncoge	1.41e-52
36	91	6.0	1604	26	HSCETS1	Human DNA for c-ets-1	1.04e-51
37	91	6.0	1604	22	196206	Sequence 43 from paten	1.04e-51
38	91	6.0	1930	21	GGCETS2	Chicken mRNA for c-ets	1.04e-51
39	92	6.0	2512	21	XLXER1A	X.laewis mRNA XEL-b fo	1.04e-51
40	92	6.0	2764	21	CHKETS1	Chicken protein p54 (e	1.41e-52
41	91	6.0	4991	29	RATERSONCA	Rat proto-oncogene (Et	1.04e-51
42	90	5.9	1762	21	GGCETS	Chicken c-ets mRNA for	7.60e-51
43	90	5.9	1856	21	GGETS	Chicken mRNA for c-ets	7.60e-51
44	89	5.8	3308	19	SUETS	Strongylocentrotus pur	5.54e-50
45	89	5.8	4032	21	XLCE1S1A	Xenopus laevis c-ets-1	5.54e-50

## ALIGNMENTS

RESULT 1	LOCUS	1516 bp	RNA	VRT	27-APR-1995
DEFINITION	G.gallus ERG mRNA.				
ACCESSION	X77159				
NID	g790439				
KEYWORDS	ERG gene.				
SOURCE	chicken.				
ORGANISM	Gallus gallus				
REFERENCE	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.				
AUTHORS	1 (bases 1 to 1516) Dhordain,P., Dewitte,F., Desplens,X., Stehelin,D. and Duterque-Cogulliaud,M.				
TITLE	Mesodermal expression of the chicken erg gene associated with precartilaginous condensation and cartilage differentiation				
JOURNAL	Mech. Dev. 50 (1), 17-28 (1995)				
MEDLINE	95329425				
REFERENCE	2 (bases 1 to 1516) Duterque-Cogulliaud,M.				
AUTHORS	Direct Submission				
TITLE	Submitted (17-JAN-1994) M. Duterque-Cogulliaud, CNRS UA 1160, Oncologie Moleculaire, Institut Pasteur, 1 rue Calmette, 59019 Lille, FRANCE				
JOURNAL					
FEATURES	Location/Qualifiers				
source	1..1516				
	/organism="Gallus gallus"				
	/db_xref="taxon:9031"				
	/dev_stage="adult"				
	/tissue_type="spleen"				
	/clone_idb="lambda gt10"				
	63..1499				
	/gene="ERG"				
	63..1499				
	/gene="ERG"				
	63..1499				
	/codon_start=1				
	/db_xref="PID:g790440"				

```

/translation="MASTIKELSVSEDSLEFCAVSPHIAKTEMTASSSEYQOT
SKMSRVRVQODMLSQPARVYIKMECNNOVNSNSNDVAGKRVSSSDNVGM
NYGSMERKIPPMPTNERNVIVADPTLNERVLEWEMAKVGLSSPDILFL
ONIDKELCKMTKDFORLESLHSHYRETPILITSDDDKALONSPRL
MHARTGATRTFPTSYSPRATORITRPDIYPOARSASMTSHSHPOKATPSS
STVRTEQORPOLDIYQILGPTSSRLANPGSQIDKOFLELSDSSNSKITTEGT
NGEEMTPDEVARWGERKSKPMNMYKLSHLYYDKNMTYVHGRVRYKRFDFH
GIAOLQHPPESSMYKYPSDLPYMSYHAHPKKNFVAHPHPALPVSSSFEAPNP
YVNSPTGGIYPTRLPAAHMPHLGTYT"
BASE COUNT      458 a      392 c      327 g      339 t
ORIGIN

```

```

Query Match      99.2%; Score 1516; DB 21; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAATTCGGGAGCAATTAATTTATTAGCAATTTATGCGATCAATTAATCTTGATCACA 60
    1 gaatttcgggagcaaatlaattatlaagcaattatagcgaatcaatcttgcacaca 60
    61 TTATGCGAGCACTATTAAAGAGCAATTTACGTGTGATGATGAAGACAGCTCTGTTG 120
    61 ttatgcgagcaactattaaagagcaatttacctgtgtgagtgagagacagctctgttgc 120
    121 AGTGTGCTACGCGATGCGCCACCTTGCAAAAGACAGAAATGACAGCCTCTCTCCAGTG 180
    121 agtgtgctacgcatgccccaccttgcaaaagacagaatgacagcctctctccagtg 180
    181 AATATGGGCAACATCAAGATGAGCCCGCGCTTCCCGACGAGACTGGTTATACAGC 240
    181 aatattgggcaaacatcaagatgagcccgcgcttcccgacgagactggttatcacagc 240
    241 CCGCGGCGAGAGTTACATTAAGATGAGTGAAGTGAACCAACAGGTTAAAGGCGTCAAGA 300
    241 cccggcgagagttacattaaagatgagtgatgaaccacaacaggttaaggggtcgaagga 300
    301 ATTACCTGATGATGCTACGCGTGGCAAAAGAGAGGAAATGTTAGCAGTTCAGACATG 360
    301 attacctgatgatgctacgctggcaaaagagaggaatggttagcagttcagacaatg 360
    361 TTGGGATGAATCTATGAGAGTACATGGAAGAGAGCAATATTCGGCTCCCAATATGCA 420
    361 ttgggatgaatctatgagagtacctggaagagagacatcttcgctcccaatatgacaa 420
    421 CCAATGAGCAAGAGTATTGTGCGACAGATCTTACGTTATGGAGCAGACACATGATAC 480
    421 ccaatgagcaagagtattgtgacagagatcttaccgattatggagcagacacatgatac 480
    481 GGCAGTGGCTGAGTGGGAGTGAAGAGTATGTTCTTCAGACGTGAGATCTTGTGT 540
    481 ggcagtggctgagtgaggagtggaagagatggttcttcagacgtgagatcttgtgt 540
    541 TCCAGAAATGATGATGGGAAAGAGTGTGTAATAATGACCAAAAGATGACTTCCAGAGACTCA 600
    541 tccagaaatgatgatgggaaagagtgtgtaataatgacccaaaagatgacttccagagactca 600
    601 CGCGAGGCTTAACGCGAGATATCTCTGTCACACCTACACTACCTAGAGAGACTCCTC 660
    601 cgcgaggcttaacgcgagatattctctgtcacacctacactcagaagagactctc 660
    661 TTCACATTTGACTTCAGATGATGTTGATAAGGCTTACAAAACCTCCACGGTGAATGCG 720
    661 ttccacatcttgacttcagatgatggttataagagcttacaacactccacaggttaatgc 720
    721 ATGCTAGAAAACACAGAGAGGCCATTTTATTTTCCAAAATACATCAGTTTACCAGAG 780
    721 atgctagaaaacacagagagggccatttattttccaaaatacatcagtttaccagag 780
    781 CAACGCAAAATATACACAAGGCGACATTTACTTTATGAGCAGGACGAGAGATCAGCCT 840
    781 caacgcaaaatatacacaagggcagatttacctttagcgaagcgagagatcagcct 840
    841 GGAGAGTCAAGCAGCATCCACTCAGTCAAAAGCTACCAACATCATCTTCAACAGTGC 900

```

```

      841 ggcagagtagcagccatccctcactcagtcacaaagctacccaacacacatcttcaacagtcg 900
      901 CAAAAGAGAAAGACAGCGCTCTCTCAGATTAGATCTTATAGATTTCTTGAGCCAGCAGCA 960
      901 ccaaaagagaaagacagcgctctcagtgtagatcctcatcagttcttgagcagcagcga 960
      961 GCCGCTCTTGCAAAATCCAGGAGTGGGAGATACAGCTATGCGAGTTCCTACTGAGCTTC 1020
      961 gccgctcttgcaaatccagagtagtgaggatcacagctatgagttcttcaactgagcttc 1020
      1021 TGTGCGACAGCTCCCAACTCCACTGATCCTGAGGAGGCGCAAAATGGGAGTCAAGA 1080
      1021 tgtgcgacagctcccaactccactgcatcaccctggagggccaaatgggaggttcaaga 1080
      1081 TGACAGACCTGATGAAGTGGCTCGCGCTGGGAGAGAGCAAAAGCAAACTTAACATGA 1140
      1081 tgacagacctgatgaagtggctcgcgctgggagagagcaaaagcaaaacttaacatga 1140
      1141 ACTATGCAAACTCAGCCGCTGCACTTCCGCTACTATGACAAAATATTAAGTAAAG 1200
      1141 actatgcaaaactcagccgctgcacttgcctactatgacaaaatatatgaactaaag 1200
      1201 TTGATGTTAAAGCTATGCTTACAAATTTGATTTCCAGGAGATCGCTCAGGCCCTCCAGC 1260
      1201 ttgatgtttaaagctatgcttacaatttgatttccaggagatcgctcaggccctccagc 1260
      1261 CTCACCTCCAGATCATGATCATGATCAATTAACCAATACCAATCAGACCTCCGCTATGAGTCT 1320
      1261 ctcacctccagatcatgatcatgatcaaatcaacatcagaacctccctcctacagttcct 1320
      1321 ACCATGACACCCGCCAAGAGATGAATTTGATGCTCCCATCCCTCCCTGTTGCCGTAA 1380
      1321 accatgacacccgccaaagatgaatttgatgctcccatccctcctgcttgcggttaa 1380
      1381 CCTGATCCAGCTTTTGTGCTGCCCCCTAATCCTACATGGAATTCACCAATGGAAGCATCT 1440
      1381 cctgatccagcttttgtgctgccccctaatcctacatcgaatcacaacatggagcatct 1440
      1441 ACCGCAATACGAGGCTGCGACGATGCTCATATGCTCCCATCTTGCGACCTACTACTAG 1500
      1441 accgcaatatcggcgtgcagcgtgcacatcagctctccatcttgagacctactaag 1500
      1501 TGGGCAAGAAAGAA 1516
      1501 tgggcaagaaagaaa 1516

RESULT 2
LOCUS      816 bp      mRNA
DEFINITION Erg-3-immunoglobulin heavy-chain enhancer-binding Ets protein
            [alternatively spliced] [mice, pre-B-cell line 22d6, mRNA Partial,
            816 nt].
ACCESSION 56169
NID       56169
KEYWORDS 9436276
SOURCE    Mus sp. pre-B-cell line 22d6.
ORGANISM Mus.
REFERENCE 1 (bases 1 to 816)
AUTHORS   Rivera,R.R., Stuijver,M.H., Steenbergen,R. and Murte,C.
TITLE     Ets proteins: new factors that regulate immunoglobulin heavy-chain
            gene expression
JOURNAL   Mol. Cell. Biol. 13 (11), 7163-7169 (1993)
MEDLINE   94019387
REMARK    Genbank staff at the National Library of Medicine created this
            entry [NCBI g1dbsg 138523] from the original journal article.
            This sequence comes from Fig. 2.
FEATURES
            1..816
            Location/Qualifiers
            /organism="Mus sp."

```

gene /db\_xref="taxon:10095"  
1. . 816  
/partial  
/note="Immunoglobulin heavy-chain enhancer-binding Ets  
protein"  
/gene="Erg-3"  
1. . 816  
/partial  
/gene="Erg-3"  
/note="Immunoglobulin heavy-chain enhancer-binding Ets  
protein. This sequence comes from Fig. 2"  
CDS  
1. . 816  
/partial  
BASE COUNT 201 a 279 c 172 g 164 t  
ORIGIN  
Query Match 36 1%; Score 552; DB 29; Length 816;  
Best Local Similarity 84.8%; Pred. No. 0.00e+00;  
Matches 691; Conservative 0; Mismatches 121; Indels 3; Gaps 2;  
Db 1 GATGATTCACGCGCTCCGCGAGCTACAAATGCGACATTCCTCTCATCTCCAC 60  
OY 582 gattgacttcagagactcagcgagactataagcagatactctctgacacccacac 641  
Db 61 TACCTCAGAGAGATCCCTTCACATCTGACTTCGATGACGTGATAGGCTTTACA 120  
OY 642 taccctagagagactcctctccacattgactcagatgattgataagccttaca 701  
Db 121 AACTCTCCAGGTTATGATGCGAAGACAGAGGGGTGACCTTTATTTCCCAAT 180  
OY 702 aactctccaggttaatgactgtagaacaagagagacacttatttcccaat 761  
Db 181 ACTGATGATATCCGAGACGCAAGAAATTAATCACTAGACGATTTACCTTAG 240  
OY 762 aatcagttacccagagacagcaagaataacaagaagccagatttactctatg 821  
Db 241 CCTCCAGAGATCAGCTGAGCGGACAGCCACCTCAGCCAGTCCAAAGTGT 300  
OY 822 caagcagagagatcagcgagagagtcacagcca--tc-cacatcagccaagctac 878  
Db 301 CAGCCATCTCCCTGCTGATGCCCAAACTGAAGACAGCAGCTCTAGTTAGTCTTAC 360  
OY 879 caacatcattctcaacagtgcccaaaacagaagaccagctcctcagttagctt 938  
Db 361 GACATCTGGAGACGACAGTAGCGCTTGTATTCAGAGTAGGCGAGATCCAGCTG 420  
OY 939 cagatcttgagacgacgacgagcgctctgcaaatccagagagtgagagatacag 998  
Db 421 TGGCAGTCTCTGCTGACACTCTGTGACAGAGCTCCCACTCACTGATCGGAA 480  
OY 999 tggcagttcctcactgagctctgtcgagacagctcactcactcactcactg 1058  
Db 481 GGCACCAAGGGGAGTTCAAGATGACAGACCGGAGAGAGTGGCTGGGGGGAG 540  
OY 1059 ggcacaaatgggagttcaagatgacagacccctgtaagtgctgagcgttgg 1118  
Db 541 AGGAAGACAGAGCCCAACATGAACTATGACAGCTCAGCGCGGCTCGCTCTACT 600  
OY 1119 aggaagaagacaaactaaactaaactaaactaaactaaactaaactaaactaa 1178  
Db 601 GACAAAACATCATGACCAAGGTGACGGGAAGCGCTACGCTTCAAGTTTGA 660  
OY 1179 gacaaaatatatagactaaagtctatggttaagcgtatgctcaacaatttg 1238  
Db 661 GGGATTGGCCAGGCGCTGAGACCCCACTCTGATGCTGCTGTTGACAGTACC 720  
OY 1239 ggaatgcctcagcgctcagcctcagcctcagaaatcattcattacaatacc 1298

Db 721 GACCTGCATACATGAGGCTCTCTATCAGCCGACCCCAAGAGATGACTTGTCTCC 780  
OY 1299 gactccctacatgagcttccactacacacccccagaagatgacttgtagctccc 1358  
Db 781 CACCTCCCGCTCTCCCACTCATCTTCCAGTTT 815  
OY 1359 catcccccctcttgcgcgtaactcactcagctt 1393  
RESULT 3  
LOCUS HUMERG2 3166 bp ss-RNA PRI 08-NOV-1994  
DEFINITION Human erg2 gene encoding erg2 protein, complete cds.  
ACCESSION M17254  
NID 9182186  
KEYWORDS erg 2 protein.  
SOURCE Human cell line COLO 320, cDNA to mRNA, clone lambda 12.  
ORGANISM Homo sapiens  
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Rao, V. N., Pappas, T. S. and Reddy, E. S.  
TITLE erg, a human ets-related gene on chromosome 21: alternative  
JOURNAL splicing, polyadenylation, and translation  
MEDLINE Science 237 (4815), 635-639 (1987)  
FEATURES  
source location/Qualifiers  
1. . 3166  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="21q22.3"  
257. . 1645  
/gene="ERG"  
257. . 1645  
/note="ERG"  
/note="erg 2 protein"  
/codon\_start=1  
/db\_xref="GDB:600-119-884"  
/translation="MIGVPPPAHIREALSVSFDLSFCAGYCTPHILATEMNAS  
SSDYGTSKMAPRPQDMLSPPARRTITKECNRSQNGSRNSDECYVAKGKRVG  
SPDIVGMNYSMEKHPPPNMTNERRVLPADPLMSTIDHVMLEMAVKEYGLP  
DVNILLFONIDGKEKMTKDDFQRLTPSYNADILSHLYLRETPPLPHLSDVDKA  
LONSRLHMHANTDLPYPPRRASATGCHPTPOSKAOPSPTVTKTEDOROLDPY  
OILGPTSSRLANPGSGOIOLOMFLLELSDSNSCTIEGNGEFTMTDDEVARRV  
GERSKPMNMTDKLSRAIRYYDKNIMKVKHGKRAKAYFEDHGICAOAOPRPRESSLY  
KPSDLPYMGSTYHAPKMFVAPHPALPVTSSEFAAPNPMYSPGILTPNTRLP  
TSHMPSHLGTY"  
repeat\_region  
/note="8-bp repeat sequence"  
291. . 298  
repeat\_region  
/note="8-bp repeat sequence"  
509. . 516  
polyA\_signal  
/note="putative"  
/note="8-bp repeat sequence"  
3143. . 3149  
BASE COUNT 227 bp upstream of PvuII site; chromosome 21.  
ORIGIN  
Query Match 35.9%; Score 549; DB 26; Length 3166;  
Best Local Similarity 81.3%; Pred. No. 0.00e+00;  
Matches 1160; Conservative 0; Mismatches 197; Indels 69; Gaps 27;  
Db 289 TATCAAGAGGCTTATGTTGTGAGTGAAGACGAGTGTGTTGAGTGCCACGG 348  
OY 74 taattaagaagatatactagtgtagtgaagacagctcctcttgtagtgctcagc 133  
Db 349 AACGCCACACTGGCTTAAGACAGAGATGACCGCTCTCTCCAGAGATGACAGAC 408  
OY 134 atcgcccccacttgcaagaagacagatgacgctcctctcactgtaatatggc 193  
Db 409 TTCGAAGATGACCCACGCGCTCCCTCAGCAGATGGCTGCTCAACCCCAAGG 468  
OY 194 atcaagaatgagccgctgctccacagcagactgtagatcaccgcccgcagag 253  
Db 469 CACCATCAAAATGAAATGTAACTTACCGAGGTGAATGGCTCAAGAACTCTCTGATGA 528

```

OY      254  taccatlaagaatgagtgtaataccaaacaggttaatggttcaaggaattaccctgtcga 313
DB      529  ATGCACTGTGGCCAAAGGCGGGAAGATGTGGGAGGCCAGCCAGCTTGGATGAACTA 588
OY      314  ctgcaagcgtggaagaaggggaatgtaagcagttcagaagaatgttggatgaacta 373
DB      589  CGGCACCTACATGAGAGGAGAACATGCCACCCCAACATGACCAAGAACGCGCAG 648
OY      374  tggaaagcttaattggaagaagcatttcgcctcaaaatgaaacaaatgaacgaag 433
DB      649  AGTTATGCTGCGCAGAGATCTACGCTATGAGATGAGATGAGATGAGATGAGATGAG 708
OY      434  agttatgtgcaagcagatctcctgaatggaacacagacagacagacagacagacag 493
DB      709  GTGGGGGTGGAAGAAATATGCGCTTCCAGAGCTCAACATCTTGTATTATCCAGACATGA 768
OY      494  gtggagcagtggaagagatgctgtctccagacgtagacattgttgcagaacattga 553
DB      769  TGGGAAGAACTGTGCAGATGACCAAGAGACGACTTCCAGAGCTCAGCCCAAGTACAA 828
OY      554  tggaaagagtgctgttaaaatgacaaatgacttccagagactcagcgcgagctaa 613
DB      829  CGCGAGCATCTCTCTGCAATCTCCACTACCTCAGAGAGACTCTCTGCAATTTGAC 888
OY      614  cgaagatattcctcgtcacacacacacacacacacacacacacacacacacacacac 673
DB      889  TTGCAATGATGTTGATTAAGCCCTTACAAACTCTCCAGGTTATGATGATGATGATGAT 948
OY      674  ttcagatgagtgatgaatgaagccttaacaaatctccacggttaatgatagtcgaacac 733
DB      949  AG-----A--TTTA-----C-----CAT-A--T-----GA-GCC-C-C----- 968
OY      734  aggaagagacacatttatttttcaaaatatacatcagtttaccacgaagcaagaat 793
DB      969  --C--GA-GG--AGATCAGGCTG--GA-C--G-G--G--TCA-CG--GGC-A--C-C-- 1001
OY      794  aaaaacaaagccagatatttacttaatgaagcagagagagacagcgtggaagagacag 853
DB      1002  CCAGGCC-C-CAGTGCAGAAAGCTGCTCAACCATCTCTTCCAGAGTGGCCCAAACTGAGA 1059
OY      854  ccaatccacacagtcacaaagctacccaacacatcatcttcaacagtgcccaaaacagaga 913
DB      1060  CCAGGCTCTCACTGATGATCTTATCATGATTTTGGACCAAGATAGCCGCTTGCAAA 1119
OY      914  ccaagctcctcagttgatcttcatcagatcttgcagacagcagcgcgtcttgaa 973
DB      1120  TCCAGGAGTGGCCAGATCCAGCTTGGAGTTCCTGCTGAGAGCTCTCTGCGAGAGCTC 1179
OY      974  tccaaggaatggaagataaagatacgaatggaagcttccacagcgtctcgcgaacagctc 1033
DB      1180  CAATCTCAGCTGATCATCTGAGGAGGAGCAACAGGAGATTCAGATGACGATCCCGA 1239
OY      1034  caactcgaactgatacaaccctggagagacaaatgagagttcaagaatgacagacctcga 1093
DB      1240  CGAGGTGGCCCGGCGCTGGGGAAGAGCGGAAGAGCAAAACCAACATGAATCAATAAGCT 1299
OY      1094  tgaugtgagcgcgcttgggagagagagaaagcaaaacccaatacgaatacctaact 1153
DB      1300  CAGCGCGGCGCTCGTACTACTATGACAAAGATCATCAGCAAGATTCATGAGGAAGCG 1359
OY      1154  cagcgtgcaactcgcgaactaactaagaacaaatataatgaactaagaatgtaagtaacg 1213
DB      1360  CTAGCGCTACAAAGTTGACTTTCCAGGGAGTGGCCAGGCGCTTCCAGCCCAAGCCCGGA 1419
OY      1214  ctatgctacaataatgttcaatcagaagatcgctcagcgccctcagcctcaacctcaga 1273
DB      1420  GTCATCTGTGTACAGTACCTCTCAGACTCCCGTACATGAGGCTCTTATACGCCCAACC 1479
OY      1274  atcatcattgataaataacccaatcagacccctccctacatgattcttaacattcagaccc 1333
DB      1480  ACAGAGATGAAGTCTTGTGGCGGCCACCTTCAGGCCCTCCCGTGAACATCTTCAAGTTT 1539

```

```

OY      1334  ccagaagatgaactttgtgctcccaatcccccctgcttggcccgtaactcatcagctt 1393
DB      1540  TTTTGTGCCCCCAACCATACTGATTAATTCACCACTGGGGGTATATACCCCAACTAG 1599
OY      1394  ttgtgctcccccttaactaactgtaattcaccaactggagcattaccaccaatccag 1453
DB      1600  GCTCCCCCAGCAGCATATGCTTCTCATCTGTGGCATTTACTATA 1645
OY      1454  gctgcagcagctcattatgcttcccatcttgcacactactaata 1499

RESULT      4
LOCUS      HUMERGL1      3126 bp      mRNA      PRI      15-MAR-1989
DEFINITION      Human erg protein (ets-related gene) mRNA, complete cds.
ACCESSION      M21535 M17990
NID      9182182
KEYWORDS      erg protein.
SEGMENT      1 of 2
SOURCE      Human, cell line COLO 320, cDNA to mRNA, lambda-7.
ORGANISM      Homo sapiens
Eukaryote; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 3126)
AUTHORS      Reddy, E.S.P., Rao, V.N. and Papas, T.S.
TITLE      The erg gene: A human gene related to the ets oncogene
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 84, 6131-6135 (1987)
MEDLINE      87317608
FEATURES
     source          location/Qualifiers
     CDS              1..3126
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        195..1286
                        /note="erg1 protein"
                        /codon_start=1
                        /translation="MVGSPDYGMVSYGMEKHPNMTNERNVYVPADPTLMST
                        DHROWLMAVKEVGLPVDNILLFQNIIDGKELKMTKDDFQRLTFSYNADILSLHT
                        LRETPRLTSDVDKALQNSRLNHNANTDLPEPPRRSAMTGHPDPSKAQPS
                        PSTVPTEDQRPDLPYOILGPTSSRLNPNDSGQIOLQFLELSDSSNSCITWEG
                        TNGEKKMDPDEVARMEGRKSRPMNNTDKLSRALRYDDKNIMTKVGRKRYAFDF
                        HGIADALPHPESSLYKRPDLPRMPSGVNHPQKMNVAHPHRLALPTSSSFFAAPN
                        PYNNSPTGIIPTNIRLPTSHMSHGLTY"
BASE COUNT      928 a 732 c 725 g 741 t
ORIGIN      1 bp upstream from EcoRI site.

Query Match      30.3%; Score 463; DB 26; Length 3126;
Best Local Similarity 84.2%; Pred. No. 0.00e+00;
Matches 588; Conservative 0; Mismatches 107; Indels 3; Gaps 1;

```

DB	949	CCOTCCGTTACTACTATGACAGAAACATCATGACCAAGGTCATCGGAAGCGCTACGCCCT	1008
OY	1162	cacttcgctactactacatgacaaatactatgactaaagtctatgataacgctatgct	1221
Db	1009	ACAAGTTGACTTGCACCGGATGCGCCAGCCCTTCACAGCCACCCCGGAGTCATCTC	1068
OY	1222	acaaattgattctacaggaatcgctcagcgccctccagccctcacccacgaatcatcca	1281
Db	1069	TGTACACTACCCCTCAGACCTCCCGTACATGSGCTCCTATACGCGCCACCCAGAGA	1128
OY	1282	tgtaaaataaccacatcagactccctccatcagatctctacatgacacacccccaaga	1341
Db	1129	TGAACTTGTGGCGCCCGACCTCCAGGCCCTCCCGGTGACATCTTCAGATTTTCTCTG	1188
OY	1342	tgaactttgtgctcccccacatccctcgcttcgctcgaatacctcatcagcttttctg	1401
Db	1189	CCCCAACCCTACTGATTCGATTCACCAACTGGGGGTATATACCCCAACTAGCGTCCCA	1248
OY	1402	ccccaaatccatctgattgattcaccaactggaagcattcaccacaaataccagctgcag	1461
Db	1249	CGAGCCATATGCTTTCATCTGCGGCACTACTACTAA	1286
OY	1462	ctgctcatactgctcccatctctgcacactactactaa	1499
RESULT	5		
LOCUS	572621	762 bp	762 bp
DEFINITION	EWS. . . erg (translocation, type 1e and 9e) [human, SK-PN-LI cell]		
ACCESSION	572621	Partial Mutant, 3 genes, 762 nt.	
KEYWORDS	g633773		
SOURCE	human SK-PN-LI cell line.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 762)		
AUTHORS	Giovannini,M., Biegel,J.A., Serra,M., Wang,J.Y., Wei,Y.H., Mrcum,L., Emanuel,B.S. and Evans,G.A.		
TITLE	EWS-erg and EWS-Fli1 fusion transcripts in Ewing's sarcoma and primitive neuroectodermal tumors with variant translocations		
JOURNAL	J. Clin. Invest. 94 (2), 489-496 (1994)		
MEDLINE	94314848		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gidsb 153609] from the original journal article. This sequence comes from Fig. 2A.		
FEATURES	Map location: t(21;22)(q22;q12).		
SOURCE	Location/Qualifiers		
	1..762	/organism="Homo sapiens"	
	/db_xref="taxon:9606"		
	1..762	/partial	
	/note="fusion gene"		
	1..85	/gene="EWS-erg"	
	/partial		
	1..762	/gene="EWS"	
	/partial		
	/gene="EWS-erg"		
	/note="This sequence comes from Fig. 2A."		
	/codon_start=1		
	/product="EWS-erg fusion protein type 1e"		
	/db_xref="PID:g633774"		
	/translation="prspptqtsysapeqiycoqssstgconlprprrrsmwtg		
	hstpskaqpsstypvkttdorpqldpfiolcptsrslanpesgdiqlmofllet		
	dsstssscitwgegnfegkmtddpdevarmererskpmnmdklsralrtyyknkn		
	vhgrrsyafkdfhgiaqlaolpdpresllykpsdldpymgsyhaahpdkmfavp		
	hpfvssfpaapnpywnspgtgilypnrflrpslshpsu"		
	86..762	/partial	
	/gene="erg"		

[illegible]

AUTHORS	Hirakawa,M., Yamaguchi,H. and Imai,K.
TITLE	Direct Submission
JOURNAL	Submitted (11-MAY-1998) to the DDBJ/EMBL/GenBank databases. Mita Hirakawa, Japan Science and Technology Corporation (JST), Advanced Database Department; 5-3, Yonban-cho, Chiyoda-Ku, Tokyo, Tokyo 102-0028, Japan (E-mail:mikahetokyo.jst.go.jp, Tel:81-3-5214-8491, Fax:81-3-5214-8470)
FEATURES	Location/Qualifiers
SOURCE	1..133746 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.2"
BASE COUNT	38362 a 29023 c 28295 g 3806 t
ORIGIN	
Query Match	23.1%; Score 353; DB 27; Length 133746;
Best Local Similarity	83.7%; Pred.No.2,06e-294;
Matches	438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
Cp	33592 TTAGTACTAGCTGCCACAGATGAGAAGCATATGGCTGTGGGGCCTACTGTTGGGTA 33651
Db	1499 ttagtagtagtgcagaatgysaaagcatalgacgtcgcaactctgatggta 1440
Cp	33652 TATACCCTCACTTGTTGTAATCCAGTAGTGGTTGGGGCGACAATAAATCGAAATGT 33711
Db	1439 gatgccctccaatcttggtgaattccagtataggattagggtggcgaaaaaagctggaatggt 1380
Cp	33712 CACGGGGAAGGGCTGAGAGGTGGGGCGCACAAAATTATCTTCGTGGTGGCGGTGATA 33771
Db	1379 taaggcgaaaagcaaggaggatgysgaagcuaaagttcatcttcctgggggtggtgatgta 1320
Cp	33772 GGAGCCCATGTACGGGAGAGTGTGAGGGCTACTTGTACAGAGATGATCCGGGGGGTGGG 33831
Db	1319 ggaactcatgtgaaggsgagtcgtatgggtatttgtacatgatatcttggaaggtgag 1260
Cp	33832 CTGGAGGGGCGTGGCGAATCCCCTGGAGACTGCAACTTTAGCGCTACAGCGCTCCATGGAC 33891
Db	1259 ctggaggggcctggagcgaatlccgtggaatacaatttgtlaagcatagtcattaccatgaac 1200
Cp	33892 CTTGTGATGATGTTCTTGTGCATAGTAGTAACGAGGCGCGGCTGAGCTTATGTAATT 33951
Db	1199 tttagtcataatatttttgtcatagtgtgycgaagtcgaagcgtggtgttgcataagtc 1140
Cp	33952 CATGTTGGGTTGCTCTTCCGCTCTCCACAGCGCGGCGCACTCGTGGGATCCGTAT 34011
Db	1139 catgttagtgttgcatttcctctctcccacaagcggagccaattcatcaggtgtcgtcat 1080
Cp	34012 CTTGAATCCCCCGTTGGTGCTTCCCAAGTGATACACTGGAGTGGAGCTGTCCGACAG 34071
Db	1079 cttyaacctccccattcttgccctcccaagttatgtaagttggaagctgtccgacag 1020
Cp	34072 GAGCTCCAGAGGAAGACTGCCAAGACTGATGTGGCACTGACT 34114
Db	1019 aaagctccagttggaactgcataagcttatctcgtccacctccct 977
RESULT	7
LOCUS	572865 588 bp mRNA PRI 24-JAN-1995
DEFINITION	EWS... EWS-erg-EWS-erg fusion protein type 9e [human, SK-PN-LI cell]
ACCESSION	line, mRNA Partial Mutant, 3 genes, 588 nt].
NID	572865
KEYWORDS	9633777
SOURCE	human SK-PN-LI cell line.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotes: mitochondrial eukaryotes; Metazoa; Chordata;
AUTHORS	Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	1 (bases 1 to 588) Giovannini,M., Biegel,J.A., Serra,M., Wang,J.Y., Wei,Y.H., Mycum,L., Emanuel,B.S. and Evans,G.A. EWS-erg and EWS-Fli1 fusion transcripts in Ewing's sarcoma and primitive neuroectodermal tumors with variant translocations

JOURNAL MEDLINE REMARK	J Clin. Invest. 94 (2), 489-496 (1994)
FEATURES	Genbank staff at the National Library of Medicine created this entry [NCBI g1b3g 154042] from the original journal article. This sequence comes from Fig. 2a.
source	location/Qualifiers 1..588
gene	/organism="Homo sapiens" /db_xref="taxon:9606" 1..588
gene	/partial /note="EMS-erg fusion protein type 9e" /gene="EMS-erg" 1..588
gene	/partial /gene="EMS" 1..588
CDS	/partial /gene="erg" 1..588 /partial /gene="EMS-erg" /note="EMS-erg fusion protein type 9e; This sequence comes from Fig. 2a." /codon_start=1 /id_xref="PID:g533778" /translation="PTSTPOTGSYSQABSOYSSQOASSGQIQIOWFLLELS DSNSSKITEGTEGFEMKTDPDEVARMRERKSPPNMNDLRLRYDDNIMTK VHKRRAYKDFHGIAOLAPHPPESSIKRYPDLVMSGYNAHQMKNFVAHPRAL PTVSSEFPAPNPYNMSPGTGIDPNTRLPSIHNSH"
BASE COUNT	145 a 210 c 129 g 104 t
ORIGIN	
Query Match	22.1%; Score 337; DB 26; Length 588;
Best Local Similarity	83.6%; Pred. No. 3.32e-279;
Matches 419; Conservative	0; Mismatches 82; Indels 0; Gaps 0;
D	8 AGTGGCAGATCCAGCTTTGGCACTTCCTCGTAGCTCCTGTGCACAGCTCCAAGTCC 147
Oy	981 agtggcagatacagctatgcagtgtccactgaagcttcgtcgacagctccaattcc 1040
D	148 AGCTGCATCACCTGGGAAGCACCACAAGGGAGATTCAAAGATGACGGATCCCGACGAGTG 207
Oy	1041 aactgcatacccttgaggagaccacaatgaggagttcaagaatgacagacctgtgaaugtg 1100
D	208 GCCCGGCTGGGGAGAGCGGAAGACAAACCCAAACTGAATGAACTGAACTAAGCTACGCCG 267
Oy	1101 gctcggcgcttgaggagagggaaaagaacctaacaatgaactatgacaaactcaagcgtc 1160
D	268 GCCCTCGTTACTACTATGACAGAACATCATGACCAAGGTCCATGGGAACGGCTACGCC 327
Oy	1161 gaactcgtactactatgacaaaaatatatgatgaacttaagtcatgtgttaaagctatgcc 1220
D	328 TACAAGTTCCAGCTTCCACGGGATCGCCCAAGCCCTCCAGCCCAACCCCCGGAGTCACT 387
Oy	1221 tacaaattcatlctccacgggaatcgcaccagggccctccaaagctccaccctccgaatctcc 1280
D	388 CTGTACAGTAAGTACCCCTGAGACCTCCGCTACATGGGTGCTCATACGCGCCACCCAGAGA 447
Oy	1281 atgtacaataaacacataagacactcccatacgttatgtctccatcacatgacacccccgaag 1340
D	448 ATGAACTTTGTGGCGCCCAACCTTCACGACCTCCCGCTGACATCTTCCAGTTTTTTTGT 507
Oy	1341 atgaactttagtattcccatccccctgctttagccgttaacctaatcagaacttlttgtc 1400
D	508 GCCCAAAACCATCTGGAATTCAACCAACTGGGGGTATATACCCCAACTAGGCTCCC 567
Oy	1401 gccctaatcatactatgaaattcaccaactggaagcatctaaccaataccaagctgcaa 1460
D	568 ACAGCATATGCTTCTCAT 588
Oy	1461 gctgtcatatgcttccat 1481



LOCUS	8	CCRNALI	3490 bp	mRNA	VRT	07-JUL-1998
DEFINITION		Coturnix coturnix mRNA for transcription factor FLI, clone 10.1.				
ACCESSION		U14773				
NID		93269302				
KEYWORDS		fil gene; FLI oncoprotein; transcription factor.				
SOURCE		quail.				
ORGANISM		Coturnix coturnix				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.				
AUTHORS		1 (bases 1 to 3490) Mager,A.M., Grapin-Botton,A., Ladfall,K., Meyer,D., Wolff,C.M., Stiegler,P., Bonnin,M.A. and Remy,P.				
TITLE		The avian fil gene is specifically expressed during embryogenesis in a subset of neural crest cells giving rise to mesenchyme				
JOURNAL		Int. J. Dev. Biol. 42 (4), 561-572 (1998)				
MEDLINE		98358003				
REFERENCE		2 (bases 1 to 3490)				
AUTHORS		Stiegler,P.				
TITLE		Direct submission				
JOURNAL		Submitted (07-SEP-1997) P. Stiegler, Centre National de la Recherche, Scientifique, U.P.R. 9005 du C.N.R.S., 15 Rue Rene Descartes, 67084 Strasbourg Cedex, FRANCE				
COMMENT		Related sequence U14774.				
FEATURES		Location/Qualifiers				
source		1..3490				
		/organism="Coturnix coturnix"				
		/db_xref="taxon:9091"				
		/chromosome="1"				
		/dev_stage="4-day quail embryo"				
		/clone_1lb="pcDNA"				
		/clone="10.1"				
		/map="q31.33"				
		215..1513				
		/gene="fil"				
		215..1513				
		/gene="fil"				
		/function="transcription factor"				
		/codon_start=1				
		/product="FLI transcription factor"				
		/db_xref="PID:e1308265"				
		/db_xref="PID:g3269303"				
		/translation="MDGTIKELALSVSDDSLPSTYGAATHLPKADMTASGNDYGC PKHINPLPPOEWINOPVRYVREKYDHMGMSRSPVDCSVNKSILVGATESPNPMS YSTMDENKGRPMNTNERVITYPADPILMTQCHYONLEMATKEYGLMEIDTFPG MMDKEICKMKKDDPLRTSLTYNEVLVLSHLYRESSLAIATPSPHTEASSRLATK EGPVAAGTQNVNKTTEQARPQDPDYQILGPTSSRLANGSQIOLMOLFELLDSSSN ASCPTWAGTNGNEFMTDPDEVARWGERKSKPNMYDKLSRALRYDYDKNIMYHVKR RYAKCFDFHGIAGLOLPHPTESMSYKPSDLSYKPSYHAHQKVNFVPPHPSMVPVTS SSFGAASPWTSFAGSIYSPAGSIYPPNPVPRHNPVSHLSYSY"				
BASE COUNT		1026 a	770 c	752 g	942 t	
ORIGIN						
Query Match		20.9%;	Score 320;	DB 21;	Length 3490;	
Best Local Similarity		79.2%;	Pred. No. 4,49e-263;			
Matches		434; Conservative	0;	Mismatches 114;	Indels 0;	Gaps 0;
Db	910	AACAAACAGCAGCAGCCCAACACAGATCCATATTAATCTCTGGGGCCACCACTAGTCG	969			
Qy	905	aacaaagaacacagcgctccagtcagtcagtcctatcagattcttgagaccgacgagcg	964			
Db	970	TCTTCACATCCTGGAGAGGGGAGATCAATCAATCTGAGAGTTCCTCGAGATTGCTGTC	1029			
Qy	965	tcttgcaaatccagggagattgggcagataacagctatggcaggtcccactggagctctgac	1024			
Db	1030	GGACAGTTCGAATGCGCACTGTATACATACATGGAGGACCAATGGGGAATTCAGATGAC	1089			
Qy	1025	ggacagctccaactccaactgcatcactgacctgggagggcacaatatgggaggtccaagtac	1084			
Db	1090	AGACCCAAATGAAATGGCGACGGCGCTGGGGGAGAAACGCAAAAGCAGCCCAATGAATTA	1149			
Qy	1085	agaccctatgaagtgagctcgcgctgtgggagagagagagagagagagagagagagagagag	1144			

LOCUS	9	CCFLIONCO	3545 bp	mRNA	VRT	07-JUL-1998
DEFINITION		Coturnix coturnix mRNA for transcription factor FLI, clone 6.1.				
ACCESSION		Y14774				
NID		93269304				
KEYWORDS		fli gene; FLI oncoprotein; transcription factor.				
SOURCE		quail				
ORGANISM		Coturnix coturnix				
		Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.				
REFERENCE		1 (bases 1 to 3545)				
AUTHORS		Mager A.M., Grafton-Bottom A., Ladjaili K., Meyer D., Wolff C.M., Stiegler P., Bonnia M.A. and Remy P.				
TITLE		The avian fli gene is specifically expressed during embryogenesis in a subset of neural crest cells giving rise to mesenchyme				
JOURNAL		Int. J. Dev. Biol. 42 (4), 561-572 (1998)				
MEDLINE		98358003				
REFERENCE		2 (bases 1 to 3545)				
AUTHORS		Stiegler P.				
TITLE		Direct Submission				
JOURNAL		Submitted (09-SEP-1997) P. Stiegler, Centre National de la Recherche Scientifique, U.P.R. 9005 du C.N.R.S. 15 Rue Rene Descartes, 67084 Strasbourg Cedex, FRANCE				
COMMENT		Related sequence Y14773.				
FEATURES		location/qualifiers				
	source	1..3545				
		/organism="Coturnix coturnix"				
		/db_xref="taxon:9091"				
		/chromosome="1"				
		/dev_stage="4-day quail embryo"				
		/clone.lib="pcdN1"				
		/clone="6.1"				
		/map="q31-33"				
		396..1595				
	gene	/gene="fli"				
		396..1595				
		/db_xref="PID:e1308263"				
		/translacion="MTASGNPDYGGPCHINDLPPOOEIMNOPVNVYKREYDHNGSR				
		ESPDCSVNCKSLKYGAGTESNPSYTYNDEKGGPPNMTNRRYIVADPIATQ				
		EHWAKIEMAIKEIGLMEIDITFPQNDGKELCKMKNKDDIERTTSLINTEYVLSHST				
		LRSSSSLLATVTPSHTEASSKLATKESPVAGTVNVKTKTEQGPQDPYQILGPTSSN				
		RLAPGSGIOIMQFLLELSDSSNACIWEGTNGEFGKTIIDPDVAVRMGERKSRKN				
	CDs					

```

misc-feature
1071..1320
/gene="fl1"
/note="ETS binding domain"
BASE COUNT      1002 a      768 g      972 t
ORIGIN

```

Query Match	Similarity	Score	Pos	Length
Best Local	79.0%	20.8%	318	535
Matches	433	Conservative	0	Mismatches 115; Indels 0; Gaps 0;

LOCUS	10	1729 bp	RNA	ROD	02-SEP-1991
DEFINITION	Mouse Fl1-1 mRNA for retroviral integration site.				
ACCESSION	X59421				
NID	950974				
KEYWORDS	c-ets-1 gene; ets gene family; fl1-1 gene; integration site.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;				
	Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;				
	Murine; Mus.				
REFERENCE	1 (bases 1 to 1729)				
AUTHORS	Letwin, K.				
TITLE	Direct Submision				
JOURNAL	Submitted (08-MAY-1991) K. Letwin, Samuel Lunenfeld Res Inst at Mount Sinai Hospital, Div of Mol and Developmental Biol, 600 University Avenue, Toronto Ontario M5G 1X5, CANADA				
REFERENCE	2 (bases 1 to 1729)				
AUTHORS	Ben-David, Y., Giddens, E.B., Letwin, K. and Bernstein, A.				

TITLE	Erythroleukemia induction by Friend murine leukemia virus: Interleukin activation of a new member of the ets gene family: Flt-1, closely linked to c-fes-1
JOURNAL	Genes Dev. 5 (6), 908-918 (1991)
MEDLINE	91257578
FEATURES	Location/Qualifiers
SOURCE	1. .1729

[illegible]

Oy	1392	ttttctgtgcacctaatccataacttggaattccaccacactgagggcactatcccccaataacc	1451
Dd	1547	AG 1548	
Oy	1452	ag 1453	
RESULT	11		
LOCUS	A36461	2916 bp	DNA
DEFINITION	Sequence 2 from Patent WO9323549.	PAT	05-MAR-1997
ACCESSION	A36461		
NID	g2293779		
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;		
AUTHORS	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae;		
TITLE	Homoe.		
JOURNAL	1 (bases 1 to 2916)		
COMMENT	Aurias,A., Delattre,O., Desmaze,C., Melot,T., Peter,M., Plougastel,B., Thomas,G. and Zucman,J. NUCLEIC ACID CORRESPONDING TO A GENE OF CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS TUMORS Patent: WO 9323549-A 2 25-NOV-1993; CENTRE NAT RECH SCIENF (FR) Other publication FR 2691475 931126 Other publication JP 85009647 960206. Location/Qualifiers		
FEATURES	1..2916		
BASE COUNT	/organism:"Homo sapiens" /db_xref:"taxon:9606" /dev_stage="ADULT"		
ORIGIN	844 a 686 c 663 g 723 t		
Query Match	20.4%; Score 312; DB 22; Length 2916;		
Best Local Similarity	78.8%; Pred.No.1,71e-255;		
Matches	427; Conservative 0; Mismatches 115; Indels 0; Gaps 0;		
Dd	882	AGAGCAACGGCCCCAGCCAGATCCGATCCGATACAGATCTTGCGCCCGACGACGATCGCTACGC	941
Oy	911	agaccagcgctccatcagttagatcctcatcagatcttcgagccgacgacgacgctctgc	970
Dd	942	CAACCTCGAAGCGGGGAGATCATCGCTGTGGCAATTCTCTCGAGCTGCTTCCACAG	1001
Oy	971	aaatccaggagctgagagatacacatgagcttctctctgagcttctgctgagacag	1030
Dd	1002	CGCCAAACGCCAGCTGATCACACTTGGAGGGGACCAACGGGAGTTCAAAATGACGGACC	1061
Oy	1031	cctccaacctccaatcgatcacctctgctgaggggagacaatcggggagcttaagaatgacagacc	1090
Dd	1062	CGATGAGGTGGCCAGCGCGCTGGGCGAGCGGAAAAAGCAAGCCCACACATGAATTACGACAA	1121
Oy	1091	tgatgaagctgcgctgcgctgagggagggagggaaaaagaacactacatagtaactatgacaa	1150
Dd	1122	GCTAGCGCGGCGCTTCGCTTAATCTGATGATAAAAACAATATATATACCAAATGACAGGGCAA	1181
Oy	1151	aactcagctgacactctgctctctctctatgaaaataatatatgactaaagttaagttaa	1210
Dd	1182	AAGATATGCTTCAAAATTGACTTCCATCCACGGCATGGCCAGGCTCTGACGACATCGGAC	1241
Oy	1211	agcctatgctcacaatatttgtttccacgagbaatgctccagggccctccagctcaactctcc	1270
Dd	1242	CGAGTCTCATGTATCAAGATACCTTGTGACATCTCTTAGATGCGCTTCTACATCGCCA	1301
Oy	1271	agaatcaccatcatgatacaataaccatcagacctccctcataagttctctaagtcagaca	1330
Dd	1302	CCAGCAGAAGGTGAAGTTTGCTCCCTCCCAACCAATCTTCATAGCTGCTTCACTTCCTCCAG	1361
Oy	1331	cccccaagaagatgaaacttgtagctcccacacccccctgcttggccgtaaacctcatccag	1390

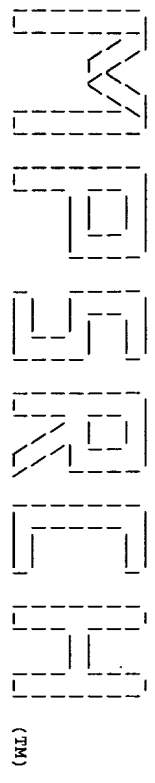
	Df	1422 CA 1423    OY 1451 ca 1452
	RESULT 12	
	LOCUS	HSHUMFLI 2938 bp RNA PRI 28-JUN-1995
	DEFINITION	H.sapiens HUMFLI-1 mRNA.
	ACCESSION	X67001 S44250
	NID	g32529
	KEYWORDS	FLI-1 gene homologue.
	SOURCE	human.
	ORGANISM	Human sapiens
	REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 2938)
	TITLE	Delattre,O.
	JOURNAL	Direct Submission Submitted (26-MAY-1992) O. Delattre, Lab. de Genet. des Tumeurs. Inst. Curie, 26 rue D'Ulm, 75231 Paris Cedex, FRANCE 2 (bases 1 to 2938) Delattre,O., Zucman,J., Pionastrel,B., Desmaza,C., Melot,T., Peter,M., Kovar,H., Joubert,I., de Jong,P., Rouleau,G., Aurias,A. and Thomas,G. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours Nature 359 (6391), 162-165 (1992)
	FEATURES	
	source	Location/Qualifiers 1..2938 /organism="Homo sapiens" /db_xref="taxon:9606" /dev_stage="adult" /tissue_type="bone marrow" /clone_idb="cDNA, Clontech HLI058" /chromosome="11q24" 143..1501 /gene="HUMFLI-1" 143..1501 /gene="HUMFLI-1" /codon_start=1 /product="homologue of the murine FLI-1 gene" /db_xref="PID:g32530" /db_xref="Swiss-Prot:Q01543" translation="MDGIRKALSVSDDSLPDASAGAAILPKADMTASGPDYGO PHNIIPPLPOEWINOPRYNRVREVDNMGNSRSPDVDCVSRSKLGGSGSNPMYV NSYMDEKNCPNPNTMRRTNRRIVPADPTLMTEDEVNQWLEMAIKESLSEIDPSFO NMGKEELCKRNKKEDFLRAITLYITEVLISLSLTLDRESSLAINTISHDOSRLSVE DPDSYVRARGAMNNMNGSLANKSPLEGAGTISKNTGEPRKMDPPDVARWGSRSKPNNYDK GSQRIQMDFLEILSDSANASCITWGTGEGFKMDPDPEVARWGSRSKPNNYDK LSLRATLYDKNKIMTKHGAKRAYAEKFHFIIAQLGOPHPRESSMYKTPDISYPSTYH AHOKVFNVFPHRPSMPVTSSFFGAASQTWSPTGSIYNPNVBRPNTHVPSSLSS YV"
	misc_feature	971..1222 /gene="HUMFLI-1" /note="ETS domain" 2390..2395 polya_signal 2411..2416 polya_signal 2405..2420 polya_signal 2908..2913 /evidence=experimental 2927
	CDS	
	gene	
	polya_site	847 a 692 c 675 g 724 t
	BASE COUNT	
	ORIGIN	
	Query Match	20.4%; Score 312; DB 26; Length 2938;
	Best Local Similarity	Pred.No.1,74e-255;
	Matches 427; Conservative	0; Mismatches 115; Indels 0; Gaps 0;





Db 1114 CGATGAGTGGCCAGGCGCTGGGGCGAGCGGAAAAAGCAAGCCCAACATGATTAAGACAA 1173  
1111 1111 111 1111 11 1111111 11 111111 11 11111  
Oy 1091 tgatgaagtggtcggcgltgggagagagaaagcaaacctaactgaactatgacaa 1150  
1174 GGTGAGCCGGCCCTCCGTTATTACTATGATATAAAACATTATGACCAAGTGACGCGCAA 1233  
11 11111 11 11 11 111111 1111 111111 1111 11 11  
Oy 1151 actcagcggtgcactcgtactactactatgacaaaaatattatgaactaaagttcatgtaa 1210  
1234 AAGATATGCTTACAAATTTGACTTCACGCGCATTTGCCAGGCTCTGACGCGCACATCCGAC 1293  
1 1 1111 11111111 111111 11 1111 11 1111 11 11  
Oy 1211 acgctaigtctacaaaatttgatttccacgaaatcgctcaagccccccagcctcacccctcc 1270  
1294 CGAGTCGTCCATGTACAAGTACCCCTCTGACATCTCTACATGCTTCCCAACATGCCCA 1353  
11 11 1111111111 1111 11 11 111111 1111 1111 11  
Oy 1271 agaataatcattacaaatacccaatcaagacctccctacatgagttcctaccatgcaca 1330  
1354 CCAGCAGAGGTGAACCTTGTCCCTCCCATCCATCCATGCTGACATCTCTCCAG 1413  
11 11111 1111111111 11111111 1 1111 11 11 11111  
Oy 1331 cccccagaagatgaacttggtaagctcccacatcccctgtcttggccgtaaccctcatccag 1390  
1414 CTCTTTTGGAGCCGCATCACAAATACTGA 1442  
11 1111 111 1 1111111  
Oy 1391 ctttttggctgcacctaatcatatcgtga 1419

Search completed: Sat Apr 10 12:38:54 1999  
Job time : 2833 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MSearch - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 7 09:43:13 1999; Maspar time 20.29 Seconds  
882,246 Million cell updates/sec

24	452	13.0	464	1	S37616	transcription factor	6.14e-64
25	450	13.0	472	1	A53236	transcription factor	1.55e-63
26	452	13.0	485	1	TVCHER	transcription factor	6.14e-64
27	447	12.9	479	1	TVCHER	transcription factor	6.22e-63
28	448	12.9	488	1	TVFVES	transcription factor	3.92e-63
29	445	12.8	468	1	TVMSER	transcription factor	1.57e-62
30	445	12.8	469	1	TVHUE2	transcription factor	1.57e-62
31	440	12.7	410	2	A45938	transcription factor	1.59e-61
32	434	12.5	623	1	S33167	transcription factor	2.53e-60
33	434	12.5	718	1	S33168	pointed protein, sp1	2.53e-60
34	425	12.3	477	1	B46396	transcription factor	1.60e-58
35	425	12.3	477	1	B46396	transcription factor	1.60e-58
36	419	12.1	510	1	S43692	transcription factor	2.53e-57
37	404	11.7	342	2	A46396	transcription factor	2.45e-54
38	407	11.7	462	1	S35534	adenovirus E1a enhanc	6.20e-55
39	407	11.7	555	1	S24061	ETS2 repressor factor	6.20e-55
40	397	11.5	548	2	S59133	ETS2 repressor factor	5.99e-53
41	395	11.4	428	1	TVHUEK	transcription factor	1.49e-52
42	395	11.4	429	2	JC4965	elk1 protein - mouse	1.49e-52
43	394	11.4	430	2	I48755	MSAP1a - mouse	2.35e-52
44	393	11.3	250	2	A54308	PE-1 protein - human	3.71e-52
45	383	11.0	453	2	B42093	serum response factor	3.51e-50

Title: >US-08-878-177-4  
Description: (1-478) from US08878177.pep  
Perfect Score: 3467  
Sequence: 1 MASTIKREALSVSEDSLFE.....IYPNRRLPAHMPSHLGTYY 478

Scoring table: PAM 150  
Gap 11

Searched: 116695 seqs, 37453910 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r58  
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 48.417; Variance 93.758; Scale 0.516

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3126	90.2	478	1	S60754	transcription factor
2	3032	87.5	486	1	TVHUEC	transcription factor
3	1776	51.2	452	1	I37565	transcription factor
4	1773	51.1	452	1	S17403	transcription factor
5	1740	50.2	453	1	S49013	transcription factor
6	1571	45.3	272	2	A54617	transcription factor
7	1401	40.4	385	1	S29844	transcription factor
8	745	22.5	173	2	A56646	transcription factor
9	691	19.9	179	2	S51226	transcription factor
10	519	15.0	103	2	S28820	transcription factor
11	467	13.5	94	2	S28820	transcription factor
12	458	13.2	268	2	S11224	transcription factor
13	458	13.2	438	1	S11225	transcription factor
14	456	13.2	440	1	A35875	transcription factor
15	456	13.2	441	1	A53988	transcription factor
16	456	13.2	441	1	TVHUE2	transcription factor
17	453	13.1	472	1	B53236	transcription factor
18	452	13.0	211	2	S24300	transcription factor
19	451	13.0	440	2	I48291	transcription factor
20	452	13.0	441	1	TVCHTE	transcription factor
21	452	13.0	454	1	A48146	GA-binding protein al
22	452	13.0	454	1	A40858	GA-binding protein al
23	451	13.0	454	2	I38739	nuclear respiratory f

ALIGNMENTS

RESULT ENTRY	1	ALIGNMENTS
TITLE	S60754	#type complete
ORGANISM	transcription factor erg - chicken	
DATE	27-Apr-1996	#formal_name Gallus gallus #common_name chicken
DATE	10-Jul-1998	#sequence_revision 23-May-1997 #text_change
ACCESSIONS	S60754	
REFERENCE	S60754	
#authors	Djordain, P.; Dewitte, F.; Desbiens, X.; Stehelin, D.;	
#journal	Mech. Dev. (1995) 50:17-28	
#title	Mesodermal expression of the chicken erg gene associated with precartilaginous condensation and cartilage differentiation.	
#cross-references	M01D:95329425	
#accession	S60754	
#status	preliminary	
#molecule_type	mRNA	
#residues	1-478	#label DHQ
#cross-references	EMBL:X77159; NID:g790439; PID:g790440	
GENETICS		
#gene	erg	
CLASSIFICATION	erg	
KEYWORDS	transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology; DNA binding; phosphoprotein; proto-oncogene; transcription factor; transforming protein	
FEATURE		
119-193	#domain ets RII regulatory region homology #label ETS2\	
312-390	#domain ets DNA-binding domain homology #label ETS	
SUMMARY	#length 478 #molecular_weight 53913 #checksum 547	
Query Match	90.2%; Score 3126; DB 1; Length 478;	
Best Local Similarity	92.1%; Pred. No. 0.00e+00;	
Matches	441; Conservative 10; Mismatches 26; Indels 2; Gaps 2;	
Db	1	MASTIKREALSVSEDSLFECAVGSPLAKTEMRASSSESGQSKMSPRPQOWLSOP 60
Qy	1	MASTIKREALSVSEDSLFECAVGSPLAKTEMRASSSESGQSKMSPRPQOWLSOP 60
Db	61	PARVTKMECPNPNQVNGSRNSPDDCSVAKGKMYSSSNMGNTGYSYEEKHIPPNNMTT 120
Qy	61	PARVTKMECPNPNQVNGSRNSPDDCSVAKGKMYSSSNMGNTGYSYEEKHIPPNNMTT 120
Db	121	NRRVIVPADPTLWSTDHVROMLEWAVEYGLPDVDILLFONIDGKELCKMTKDDFORLT 180
Qy	121	NRRVIVPADPTLWSTDHVROMLEWAVEYGLPDVDILLFONIDGKELCKMTKDDFORLT 180



```

Db      181 PSYNADILLSHLYLLETPEPHLITSDVDVALONSPRLMARNTGATFTFPNTSVPEA 240
      181 PPSYNADILLSHLYLLETPEPHLITSDVDVALONSPRLMARNTGATFTFPNTSVPEA 240
Oy      181 PSYNADILLSHLYLLETPEPHLITSDVDVALONSPRLMARNTGATFTFPNTSVPEA 240
Db      241 TORITTRPDLPEQARRSAMTSHSHPTOSKATOPSSSTVTKTEDQRPDLDPYQILGPTSS 300
      241 TORITTRPDLPEQARRSAMTSHSHPTOSKATOPSSSTVTKTEDQRPDLDPYQILGPTSS 300
Oy      241 TORITTRPDLPEQARRSAMTSHSHPTOSKATOPSSSTVTKTEDQRPDLDPYQILGPTSS 300
Db      301 RLANPGSGGIGLOMQLLELLSDSSNSNCITWEGTNGEFKKTDPDEVARNRGERSKRNMN 360
      301 RLANPGSGGIGLOMQLLELLSDSSNSNCITWEGTNGEFKKTDPDEVARNRGERSKRNMN 360
Oy      301 RLANPGSGGIGLOMQLLELLSDSSNSNCITWEGTNGEFKKTDPDEVARNRGERSKRNMN 360
Db      361 YKLSRALYUYDDKNTMTKVGKRYAVKPDFHGIQAOLQCHPRESSMYKYP-SDLPWSS 419
      361 YKLSRALYUYDDKNTMTKVGKRYAVKPDFHGIQAOLQCHPRESSMYKYP-SDLPWSS 419
Oy      361 YKLSRALYUYDDKNTMTKVGKRYAVKPDFHGIQAOLQCHPRESSMYKYP-SDLPWSS 419
Db      420 YHAPDKMNFVAPHPRALPYTSSSFPAAPNPYNSPTGGIYPTRLPAAHMPSHLGTYY 478
      420 YHAPDKMNFVAPHPRALPYTSSSFPAAPNPYNSPTGGIYPTRLPAAHMPSHLGTYY 478
Oy      420 YHAPDKMNFVAPHPRALPYTSSSFPAAPNPYNSPTGGIYPTRLPAAHMPSHLGTYY 478
RESULT 2
ENTRY TVHUEG #type complete
TITLE transforming protein erg-3 - human
CONTAINS transforming protein erg-1: transforming protein erg-2
ORGANISM #organism Homo sapiens #common_name man
DATE 31-Mar-1989 #sequence_revision 30-May-1997 #text_change 31-Oct-1997
ACCESSIONS A94294; A94178; I58410; A28041; A29515
REFERENCE A94294
#authors Rao, V.N.; Papas, T.S.; Shyam, E.; Reddy, P.
#journal Science (1987) 237:635-639
#title erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation.
#cross-references MUID:87263429
#accession A94294
#molecule_type mRNA
#residues 1-231,256-486 ##label REA
#cross-references GB:M1254; NID:g182187; PID:g182187
REFERENCE A94178
#authors Reddy, E.S.P.; Rao, V.N.; Papas, T.S.
#journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:6131-6135
#title The erg gene: a human gene related to the ets oncogene.
#cross-references MUID:87317608
#accession A94178
#molecule_type mRNA
#residues 100-231,256-486 ##label REB
#cross-references GB:M1535; NID:g182182; PID:g182185
REFERENCE I58410
#authors Prasad, D.D.; Rao, V.N.; Lee, L.; Reddy, E.S.
#journal Oncogene (1994) 9:669-673
#title Differentially spliced erg-3 product functions as a transcriptional activator.
#cross-references MUID:94119611
#accession I58410
#status preliminary: translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 230-259 ##label RES
#cross-references GB:S68130; NID:g544960; PID:g544961
GENETICS GDB:ERG
#gene GDB:ERG
#cross-references GDB:119884; OMIM:165080
#map_position 21q22.2-21q22.2
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain
#homology: ets RII regulatory region homology
KEYWORDS alternative splicing; DNA binding; nucleus; proto-oncogene;
transcription factor; transforming protein
FEATURE 1-231,256-486
#product transforming protein erg-2 #status predicted
#label EG2\
100-231,256-486 #product transforming protein erg-1 #status predicted
#label EG1\
126-200 #domain ets RII regulatory region homology #label ETS2\

```

Query Match	87.5%	Score 3032	DB 1	Length 486
Best Local Similarity	87.5%	Pred. No. 0.00e+00		
Matches 419	Conservative	28	Mismatches 29	Indels 3
			Gaps 3	
Db	9	AAHKELAVSEDSQSLFECAYGTGPLATKEMTASSSSDYGTSKMSRVPQOWLSQPP	68	
Qy	2	ASTKEALVSVSEDSQSLFECAYGTGPLATKEMTASSSSDYGTSKMSRVPQOWLSQPP	61	
Db	69	ARVATIKKECNPQVNGSRNSPDECSYAKGKMGVSPDTQGNMGYSYMEKKHPPNMTTN	128	
Qy	62	ARVATIKKECNPQVNGSRNSPDECSYAKGKMGVSSDNGMGYSYMEKKHPPNMTTN	121	
Db	129	ERRVIVPADPTLWSTDHVRQMLEMAVKEYGLPDVNIILFQNIIDGKELCKMTKDFQRLTP	188	
Qy	122	ERRVIVPADPTLWSTDHVRQMLEMAVKEYGLPDVNIILFQNIIDGKELCKMTKDFQRLTP	181	
Db	189	SYNDDILLSHHTYREPLPHLTSDDVDKALONSPRLHMAKRTGAAATFFPRTSYVPAT	248	
Qy	182	SYNDDILLSHHTYREPLPHLTSDDVDKALONSPRLHMAKRTGAAATFFPRTSYVPAT	241	
Db	249	QRTTRPDLPEPPRSFASMTGHHPTPOSKAASPSTYPTKEDRPOLDLPQILGPPSS	308	
Qy	242	QRTTRPDLPEPPRSFASMTGHHPTPOSKAASPSTYPTKEDRPOLDLPQILGPPSS	300	
Db	309	RLANPGSGQIQIQWQFLLELLSDSSNSCITWEGTNGEFTKMDPDEVARWGRSKPMN	368	
Qy	301	RLANPGSGQIQIQWQFLLELLSDSSNSCITWEGTNGEFTKMDPDEVARWGRSKPMN	360	
Db	369	YDKLSRALRYTYDKNIMTKYHGRRYAKKDFHGIQAOLQPHPESSLYKPR-SDLPIYGS	427	
Qy	361	YDKLSRALRYTYDKNIMTKYHGRRYAKKDFHGIQAOLQPHPESSLYKPR-SDLPIYGS	419	
Db	428	YHAPOKMNFPVAPRPAALPTSSSFPAAPNPVNSPTGTYNTRFLPSHPSHGTGY	486	
Qy	420	PHAPQKMNFPVAPRPAALPTSSSFPAAPNPVNSPTGTYNTRFLPSHPSHGTGY	478	
RESULT	3			
ENTRY	137565	#type complete		
TITLE		transforming protein fl1, long splice form - human		
ALTERNATE_NAMES		friend leukemia integration protein 1; transcription factor ERG		
ORGANISM		#formal_name Homo sapiens #common_name man		
DATE	04-Oct-1996	#sequence_revision 30-May-1997	#text_change	
ACCESSIONS	137565	S29843	S53506	A49000
REFERENCE	137378		A49015	I54170
#authors	DeLattre, O.; Zucman, J.; Plougastel, B.; Desmaze, C.; Melot, T.; Peter, M.; Kovar, H.; Joubert, I.; de Jong, P.; Rouleau, G.; Aurias, A.; Thomas, G.			
#journal	Nature (1992) 359:155-165			
#title	Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours.			
#cross-references	MUID:92396239			
#accession	I37565			
#status	preliminary; translated from GB/EMBL/DBJ			
#molecule_type	mRNA			
#residues	1-452	#label RES		
#cross-references	EMBL:X67001; NID:g32529; PIR:g32530			
REFERENCE	S29843			
#authors	Homas, R.; May, W.; Denny, C.; Raskind, W.; Moore, J.; Makl, R.A.; Beck, E.; Klemenz, M.J.			
#journal	Biochim. Biophys. Acta (1993) 1172:155-158			
#title	Human FLI-1 localizes to chromosome 11Q24 and has an aberrant transcript in neuroepithelioma.			
#cross-references	MUID:9317679			
#accession	S29843			
#molecule_type	mRNA			
#residues	1-68	'V', '70-129', 'A', '131-132', 'V', '134-322', 'Q', '324-425', '427-452' #label HRO1		



Oy	1	MASTIKREALSVSDOSLEFCAYGSP-HLAKTETMTASSSSEFYQTSTKMSRPVYQOQMLSO	59
Db	61	P-VRYNVKKEYP--HANGSRESPVDCSVSKCNKLVGCGEANGPNNYSYMDENKNGPPPPNM	117
Oy	60	PARVATIKMECHPNQNGNSRNSPDOSVAKGKGKVVSSDDVGNKNGSYMEKH-IPPPNM	118
Db	118	TTNERRVYVPADPTLMTQCHVROMLEMAIKEGGLMIDTISFFONMDGKELCKNNKEDFLR	177
Oy	119	TTNERRVYVPADPTLMTSTJHVROMLEMAKEYGLPVPDILFONIDGKELCKMTKIDFOR	178
Db	178	ATSAVTEVLEHLSLYRESS--LAVYNTTSHN-DOSSRL-NVKE--D---P--S-Y-222	
Oy	179	LTPSTINADLLSHLYKRETPPLHLLTSDVDKALQNSPRLMHAKRNGGATFIRPNTSVYR	238
Db	223	DSV-R---R-G-AMNNMNSG-LNKS-PLUG-GSQ----TMGKTEQRPQDPYQILGPT	269
Oy	239	EATQRIITTRPDLPEEGARRSAMTSHSHPRQSKATQSSSVPRTEQDQRPDLQYILGPT	298
Db	270	SGRLANPGSGQIQLOWFLELLSDSANASCITWEGTNGEFKMTDPDEVARHGERSKSPN	339
Oy	299	SSRLANPGSGQIQLOWFLELLSDSSNSNCITWEGTNGEFKMTDPDEVARMEERKSKPN	358
Db	330	MYDLSRLRLRYYYDKNIMTKYNGKRAYAKFDFHGAQAQAQPHRTETSMKKYF-SDTSYM	368
Oy	359	MYDLSRLRLRYYYDKNIMTKYHPPSS--MYKPSDLPYSSHYGKRAYKDFHGAQA	417
Db	389	PSYHAHQKVNVPSPHSSMPVTSSEFFGAASQVWTSPTAGTIPNPSVPRHPTVPSHL	448
Oy	418	LQPHAHPOKMNVTAPRPPALPVYSSSEFFAAPNRYMNSPFIGITP-N--RLPAHMHSHL	474
Db	449	GSY 452	
Oy	475	CTY 478	
RESULT	5		
ENTRY		S49013	#type complete
TITLE			transforming protein fl1 - African clawed frog
ALTERNATE_NAMES			Friend leukemia integration protein 1; transcription factor ERGB
ORGANISM			#formal_name Xenopus laevis #common_name African clawed frog
DATE		07-May-1995	#sequence_revision 23-May-1997 #text-change 05-Sep-1997
ACCESSIONS		S49013	
REFERENCE		S49013	
authors			Meyer, D.; Wolff, C.M.; Stiegler, P.; Senan, F.; Befort, N.; Befort, J.J.; Remy, P.
#journal			Mech. Dev. (1993) 44:109-121
#cite			Xl-fl1, the Xenopus homologue of the fl1-1 gene, is expressed during embryogenesis in a restricted pattern evocative of neural crest cell distribution.
#accession		S49013	
##status			preliminary
##molecule_type		mRNA	
##residues		1-453	##label MEY
##cross-references		EMBL:X66979; NID:g505486; PID:g505487	
GENETICS			
CLASSIFICATION			fl1
KEYWORDS			#superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology DNA binding; nucleus; proto-oncogene; transcription factor; transforming protein
FEATURE			
117-191			#domain ets RII regulatory region homology #label ETS2\
284-362			#domain ets DNA-binding domain homology #label ETS
SUMMARY			#length 453 #molecular-weight 51015 #checksum 774
Query Match			50.2%; Score 1740; DB 1; Length 453;
Best Local Similarity			56.4%; Pred. No. 0.00e+00;
Matches			273; Conservative 91; Mismatches 83; Indels 37; Gaps 25;
Db	1	MDGTIKREALSVSDOSLFDASVAGASHLSKADMTASANDYQOPKINDIPPOQDWINQ	60

QY	1	MASTIKREALSVSEDOSLFECAAY-SPLAKTDMTASSSSEFGQTSKMSPRVQOQMSLO	59
Db	61	P-MRVNIRKEEYE--HNNGSRESVPDSCINKCSKLIGSGSGNANTY--TYDENKNGPPPPN	116
QY	60	PARATYIKNECPNQNQNGSRNSPDSVAKGKGMWSSSDNVGNNGSYMEKEH-IPPPNN	118
Db	117	TNNRRVRYPADPALMSODHVROWLEMAIKEGVEIDCSLFONIDGKELCKSKSEDFLR	176
QY	119	TNNRRVRYPADPALMTSTHVRQWLEMAKEGVEIDCSLFONIDGKELCKSKTKDDFOR	178
Db	177	STSIYNEVLLSHLVNLRDSS--SSIGYNQAHN--DOSSRLT-AKE--D----P--S-Y-222	
QY	179	LTPSTINADLLSHLHTRKPLPHLTSDDVKALQNSPRMHRKNGCATFIRPNTSVYR	238
Db	223	EAV-R---RSG--WGNMSSP--VTKSPM-G-GTQ--NVNK-S-GDOQRSPDPDYQILGPT	270
QY	239	EAQTORTTTPDLPEEDARRSANTSHHPQSKATQSSSTVPTEDQRPQLDPTQILGPT	298
Db	271	SSRLANPGSGGIQIQLWQFLLELLSDSSNASCTWEGTNGEFKMTDPDEVARRMGERKSPN	330
QY	299	SSRLANPGSGGIQIQLWQFLLELLSDSSNSNCITWEGTNGEFKMTDPDEVARRMGERKSPN	358
Db	331	NMYDLSRLALRYKYKSKITKYGKRYAKKDPDHGIAQLQHPRTTSMKKY-PSEFSYM	389
QY	359	NMYDLSRLALRYKYKSKITKYGKRYAKKDPDHGIAQLQHPRTTSMKKY-PSEFSYM	417
Db	390	PSYSHQOQVNVNVPSPHSPSPVTSQGFATSPYNSPANIYPNPNVPHPTVHQSLL	449
QY	418	LDPHAIHPQKMNVAHPHPALPVTSSSFPAAPNRYNMSPTGIGTIPN---RLPAHHPSHL	474
Db	450	GGFY 453	
QY	475	GTYY 478	
RESULT	6		
ENTRY		A54617 #type fragment	
TITLE		transcription factor erg - mouse (fragment)	
ALTERNATE_NAMES		Ig heavy chain enhancer-binding protein	
ORGANISM		#normal_name Mus musculus #common_name house mouse	
DATE		25-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997	
ACCESSIONS		A54617	
REFERENCE		A54617	
#authors		Rivera, R.R.; Stulver, M.H.; Steenbergen, R.; Murro, C.	
#journal		Mol. Cell. Biol. (1993) 13:7163-7169	
#title		Ets proteins: new factors that regulate immunoglobulin heavy-chain gene expression.	
#cross-references		MUID:94019387	
#status		A54617	
##molecule_type		preliminary	
##residues		1-272 #label RIV	
##experimental_source		pre-B-cell line 22D6	
##note		sequence extracted from NCBI backbone (NCBIN:138523, NCBI:138524)	
CLASSIFICATION		#superfamily transcription factor erg; ets DNA-binding domain	
KEYWORDS		human; ets RII regulatory region homology	
FEATURE		DNA binding; phosphoprotein; proto-oncogene; transcription factor; transforming protein	
140-218		#domain ets DNA-binding domain homology #label ETS	
SUMMARY		#length 272 #checksum 3754	
Query Match		45.3%; Score 1571; DB 2; Length 272;	
Best Local Similarity		82.4%; Pied. No. 6, 02e-302;	
Matches		225; Conservative 17; Mismatches 28; Indels 3; Gaps 3;	
Db	1	DDFORLTPSYNDIILSHLYREPLPLTSDVDKALQNSRLMHRNTGAATFIPN	60
QY	174	DDFORLTPSYNDIILSHLYREPLPLTSDVDKALQNSRLMHRNTGAATFIPN	233
Db	61	TSVYEATQRTIRPDLPEPPRRSAMTSHSLTPQSKAQPSPAVPTEDQRPOLDPY	120

||||| 234 TSYPEATRITTRPDLPEOARRSAMTSHSHT -OSKATOPSSSTVPKTEDORPOLDDY 292  
Db 121 QILGPSSRLANGSGOIQLOMFLLELLSDSSNCTITWGTGGEKMTDPDVARME 180  
Qy 293 QILGPSSRLANGSGOIQLOMFLLELLSDSSNCTITWGTGGEKMTDPDVARME 352  
Db 181 RSKSPMANKDKSLRALRYYYDKIMTKVHGKRAYKDFEFGIALQAPHPPESSLYKYP- 239  
Qy 353 RSKSPMANKDKSLRALRYYYDKIMTKVHGKRAYKDFEFGIALQAPHPPESSLYKYP- 411  
Db 240 SDLPYGSYHAPQKMYSPHPALPVTSSSF 272  
Qy 412 HGIAQALQPHAHPOKMFVAPHPALPVTSSSF 444

RESULT 7  
ENTRY S29844 #type complete  
TITLE transforming protein fl1, short splice form - human  
ALTERNATE\_NAMES Friend leukemia integration protein 1; transcription factor  
ORGANISM ERGB  
DATE #formal\_name Homo sapiens #common\_name man  
02-Dec-1993 #sequence\_revision 23-May-1997 #text\_change  
ACCESSIONS S29844  
REFERENCE S29843  
#authors Hromas, R.; May, W.; Denny, C.; Raskind, W.; Moore, J.; Maki,  
#journal R.A.; Beck, E.; Klemasz, M.J.  
#title Blochm. Biophys. Acta (1993) 1172:155-158  
Human FLI-1 localizes to chromosome 11Q24 and has an aberrant  
transcript in neuroepithelioma.  
#cross-references MUID:93176799  
#accession S29844  
#status preliminary; nucleic acid sequence not shown;  
translation not shown  
#molecule\_type mRNA  
#residues 1-385 #label HRO  
#cross-references EMBL:M93255; NID:q18265; PID:q182661  
#note the nucleotide sequence was submitted to the EMBL Data  
Library, May 1992

GENETICS  
#gene GDB:FLI1  
#cross-references GDB:127565; OMIM:193067  
#map\_position 11q24.1-11q24.3  
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain  
homology; ets RII regulatory region homology  
KEYWORDS alternative splicing; DNA binding; nucleus; proto-oncogene;  
transcription factor; transforming protein

ATURE  
#domain ets RII regulatory region homology #label ETS2\ 52-126  
#domain ets DNA-binding domain homology #label ETS 217-295  
#length 385 #molecular\_weight 43437 #checksum 4889

Query Match 40.4%; Score 1401; DB 1; Length 385;  
Best Local Similarity 56.4%; Pred. No. 5.17e-265;  
Matches 229; Conservative 69; Mismatches 73; Indels 35; Gaps 21;

Db 10 ARESPTDCSVKSKLYGGESIPMNYNSYMDKNGPPPPNMTTNRRTVVPADATLVQ 69  
Qy 78 SRNSPDDCSVAKGKGVSSSDNYGMYSMEKH-IPPNMTTNRRTVVPADATLVQ 136  
Db 70 EHVROWLENAIKESYLMEDTSFQOMMDGKCKMKDEFLRTTLYNTEVLLSHSYLR 129  
Qy 137 DHVROWLENAIKESYLMEDTSFQOMMDGKCKMKDEFLRTTLYNTEVLLSHSYLR 196  
Db 130 ESSLP--LAYNTSHT--DOSSRL-SYKE--D---P--S-Y-DSV-R--R-G-ANGNNN 168  
Qy 197 ETPPLRLTSDDDVDKALQNSPRLMARNATGATTFIPNTSVYPRATGRTTRPDLPEOAR 256  
Db 169 NSG-LNKS--PLUG-GAO---TISKNTGCRPPDPYQILGPSSRLANGSGOIQLOMFL 221  
Qy 257 RSMATSHSHPTOSKATOPSSSTVPKTEDORPOLDDYQILGPSSRLANGSGOIQLOMFL 316

Db 222 LELSDSANASCTITWGTGGEKMTDPDDEVARRMGKSPMNNYDKLSRALRYYYDKNI 281  
Qy 317 LELSDSANASCTITWGTGGEKMTDPDDEVARRMGKSPMNNYDKLSRALRYYYDKNI 376  
Db 282 MKVHGKRAYKDFEFGIALQAPHPPESSMYKYP-SDISYMPSYAHQOKNFVPPHS 340  
Qy 377 MKVHGKRAYKDFEFGIALQAPHPPESSMYKYP-SDISYMPSYAHQOKNFVPPHS 435  
Db 341 SMPVTSSSFEGAASQWTS-TGDIYPNRPVPRPHNTHVSHSGSY 385  
Qy 436 ALPVTSSSFEFAAPNPYNSPTGTGTPNT--RLLPAHMPSHLGTYY 478

RESULT 8  
ENTRY A56646 #type fragment  
TITLE transcription factor erg/Flt-1 homolog - sea urchin  
ALTERNATE\_NAMES (Lytechinus variegatus) (fragment)  
ORGANISM #formal\_name Lytechinus variegatus #common\_name variegated  
DATE 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change  
08-Sep-1997  
ACCESSIONS A56646  
REFERENCE A56646  
#authors Qi, S.; Chen, Z.Q.; Papas, T.S.; Lautenberger, J.A.  
#journal DNA Seq. (1992) 3:127-130  
#title The sea urchin erg homolog defines a highly conserved  
erg-specific domain.  
#accession A56646  
#status preliminary  
#molecule\_type DNA  
#residues 1-173 #label Q1A  
#cross-references GB:M1067; NID:q161310; PID:q161311  
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain  
homology; ets RII regulatory region homology  
KEYWORDS DNA binding; nucleus; transcription factor  
FEATURE 6-84  
SUMMARY #domain ets DNA-binding domain homology #label ETS  
#length 173 #checksum 5090

Query Match 21.5%; Score 745; DB 2; Length 173;  
Best Local Similarity 61.8%; Pred. No. 1.85e-124;  
Matches 110; Conservative 30; Mismatches 26; Indels 12; Gaps 11;

Db 1 SGOIQLOMFLLELLSDSSNANCTITWGTGGEKMTDPDVARMGKSPMNNYDKLSR 60  
Qy 307 SGOIQLOMFLLELLSDSSNANCTITWGTGGEKMTDPDVARMGKSPMNNYDKLSR 366  
Db 61 ALRYYYDKIMTKVHGKRAY-KFDFAGLAQAMPVQADPSMYRYO-SDLT-VLPGY-HP 116  
Qy 367 ALRYYYDKIMTKVHGKRAY-KFDFAGLAQAMPVQADPSMYRYO-SDLT-VLPGY-HP 424  
Db 117 TKLNFYGT-PIN-PTINASLFSHSSSYWSSPTGANIYPSGHVTHPHASHMSSHLGTYY 172  
Qy 425 OKKNFVAPHPALPVTSSSFEFAAPNPYNSPTGTGTPNT--RLLPAHMPSHLGTYY 478

RESULT 9  
ENTRY S51226 #type fragment  
TITLE transcription factor erg/Flt-1 homolog - polychaete (Nereis  
diversicolor) (fragment)  
ORGANISM #formal\_name Nereis diversicolor #common\_name sandworm  
DATE 15-Jul-1995 #sequence\_revision 23-May-1997 #text\_change  
30-May-1997  
ACCESSIONS S51226  
REFERENCE S51226  
#authors Lelievre-Chocteau, A.; Laudet, V.; Flourens, A.; Begue, A.;  
#journal Leponce, D.; Fontaine, F.  
#title FEBS Lett. (1994) 354:62-66  
#accession Identification of two ets related genes in a marine worm, the  
#status polychaete annelid Nereis diversicolor.  
#molecule\_type DNA

CLASSIFICATION	##residues 1-179 ##label LEL	##superfamily transcription factor erg; ets DNA-binding domain
KEYWORDS	homology; ets RII regulatory region homology	
FEATURE	DNA binding; nucleus; transcription factor	
7-85		
SUMMARY	##domain ets DNA-binding domain homology #label ETS	
	length 179 #checksum 4104	
Query Match	19.9%; Score 691; DB 2; Length 179;	
Best Local Similarity	57.5%; Prcd. No. 3,93e-113;	
Matches	104; Conservative 36; Mismatches 28; Indels 13; Gaps 12;	
Db	1 GSGGQIQLOMFLLELSDSSNSHIITTEGTGCEKLVDPDEARMRGKRKSPNNNYDKLS 60	
Oy	306 GSGGQIQLOMFLLELSDSSNSCITTEGTGCEKMTDPDEVARMRGKRKSPNNNYDKLS 365	
Db	61 RALRYRYDKNIMTKVHGKRYA-YKFDPAQLAQMOPSTTDPAAYKXQ-QDILMS-GYH-H 116	
Oy	366 RALRYRYDKNIMTKVHPPESSMYKY-SDLR-YMSSYHGKRYAYKFDPAQLAQLPHAH 423	
Oy	117 TSKLMLMAAHAP-MASSASGCFPPPPAPYMSLVGSNLYPISNHSNHHGCHSSHLGSY 175	
Oy	424 POKMNFVAPHAPALPVTSSEFFAAPRPYMSPTGG-IYPN-T-R-LP-AAHMPSHLGTY 477	
Db	176 Y 176	
Oy	478 Y 478	
RESULT	10	
ENTRY	S28823	#type fragment
TITLE	transcription factor ets-21C - fruit fly (Drosophila melanogaster) (fragment)	
ALTERNATE_NAMES	transforming protein ets-6	
ORGANISM	#formal_name Drosophila melanogaster	
DATE	31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997	
ACCESSIONS	S28823	
REFERENCE	S28819	
#authors	Chen, T.; Bunting, M.; Karim, F.D.; Thummel, C.S.	
#journal	Dev. Biol. (1992) 151:176-191	
#title	Isolation and characterization of five Drosophila genes that encode an ets-related DNA binding domain.	
#accession	S28823	
#molecule_type	DNA	
##residues	1-103 ##label CHE	
##cross-references	EMBL:M88475	
GENETICS		
#gene	Ets21C	
##cross-references	Flybase:FBgn0005660	
CLASSIFICATION	#introns 64/2	
KEYWORDS	##superfamily transcription factor erg; ets DNA-binding domain	
FEATURE	homology; ets RII regulatory region homology	
8-86	DNA binding; nucleus; transcription factor	
SUMMARY	##domain ets DNA-binding domain homology #label ETS	
	length 103 #checksum 5825	
Query Match	15.0%; Score 519; DB 2; Length 103;	
Best Local Similarity	77.6%; Prcd. No. 1.57e-77;	
Matches	66; Conservative 11; Mismatches 7; Indels 1; Gaps 1;	
Db	2 GSGGQIQLOMFLLELSDSSNANASISWEGSGEFLIDPDEVARMRGKRKSPNNNYDKLS 61	
Oy	306 GSGGQIQLOMFLLELSDSSNSCITTEGTGCEKMTDPDEVARMRGKRKSPNNNYDKLS 365	
Db	62 RALRYRYDKNIMTKVHGKRYA-YKF 85	
Oy	366 RALRYRYDKNIMTKVHPPESSMYKY 390	
RESULT	11	
ENTRY	S28820	#type fragment

DB	ENTRY	RESULT	ENTRY	RESULT
	TITLE	12	S11224	#type fragment
	ALTERNATE_NAMES		transcription factor ets-1b - African clawed frog (fragment)	
	ORGANISM		formal_name Xenopus laevis #common_name African clawed frog	
	DATE		21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 08-Sep-1997	
	ACCESSIONS		S11224	
	REFERENCE		S11224	
	#authors		Stiegler, P.; Wolff, C.M.; Balczinger, M.; Hitzelin, J.; Seman, F.; Meyer, D.; Ghysdael, J.; Stohelin, D.; Bofort, N.; Remy, P.	
	#journal		Nucleic Acids Res. (1990) 18:5298	
	#title		Characterization of Xenopus laevis cDNA clones of the c-ets-1 proto-oncogene.	
	#cross-references XOTD:90384849			
	#accession		S11224	
	##molecule_type mRNA			
	##residues		1-268 #label STR	
	##cross-references EMBL:X52691; NID:g64616; PID:g64617			
	CLASSIFICATION		#superfamily transcription factor ets; ets DNA-binding domain	
	KEYWORDS		homology; ets RII regulatory region homology	
	FEATURE		DNA binding; nucleus; phosphoprotein; transcription factor	
	164-242			
	SUMMARY		#domain ets DNA-binding domain homology #label ETS	
	Query Match		#length 268 #checksum 3247	
	Best local Similarity		13.2%; Score 458; DB 2; Length 268;	
	Matches		60; Conservative 25; Mismatches 30; Indels 1; Gaps 1;	
137	DRAEINKDKVITAAALAGTGGSPQLQWQFLLELLTKSCSFISWTGDMERKISPD	196		
285	QRQLDQVQLTGTSRLNPNPGSQQLQWQFLLELLSDSSNNCITWGTGGERKMDPD	344		

#title	Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus.
#cross-references	MolUD:90299137
#accession	A35875
##molecule_type	mRNA
##residues	170-440
##cross-references	EMBL:X53953
GENETICS	
#gene	Ets1
#map_position	9 15.0
CLASSIFICATION	#superfamily transcription factor ets; ets DNA-binding domain homology; ets RII regulatory region homology alternative splicing; DNA binding; nucleus; phosphoprotein; proto-oncogene
KEYWORDS	
FEATURE	
57-130	#domain ets RII regulatory region homology #label ETS2\
97-130	#region helix-loop-helix #status predicted\
337-415	#domain ets DNA-binding domain homology #label ETS\
377-383	#region nuclear location signal
SUMMARY	#length 441 #molecular-weight 50422 #checksum 6268
Query Match	13.2% Score 456; DB 1; Length 441; Best Local Similarity 51.7%; Pred. No. 9,606-65;
Matches	60; Conservative 25; Mismatches 30; Indels 1; Gaps 1;
Db	310 DRADLNKDRPYIPAAALAGYTGSGPIQLMQLLELLTDKSCQSFISWTGDGEFKLSPD 369
Oy	285 QRPQLDPQIIGPITSSRLANPQSGIQLMQLLELLSDSSNSNCITGCTNGCFKMTDP 344
Db	370 EVARWRGKRKNPKMNYEKLRSGLRYRYDKNIHKTKAGKRYVRFVCDLQSLG 424
Oy	345 EVARWRGKRKSPKNNYNDKLSRALRYRYDKNIMTKVHPPESSMYKYPDPDLYMSY 400
RESULT	15
ENTRY	A53988
TITLE	#type complete
ALTERNATE_NAMES	transcription factor ets-1, splice form a - rat transcription factor p54; transforming protein ets-1a; transforming protein Tpl-1
ORGANISM	#formal name Rattus norvegicus #common name Norway rat
DATE	09-Oct-1994 #sequence_revision 23-May-1997 #text_change 23-May-1997
ACCESSIONS	A53988
REFERENCE	A53988
#authors	Bellacosa, A.; Datta, K.; Bear, S.E.; Patrioticis, C.; Iazo, P.A.; Copeland, N.G.; Jenkins, N.A.; Tschlis, P.N.
#journal	J. Virol. (1994) 68:2320-2330
#title	Effects of provirus integration in the Tpl-1/Ets-1 locus in Moloney murine leukemia virus-induced rat T-cell lymphomas: levels of expression, polyadenylation, transcriptional initiation, and differential splicing of the Ets-1 mRNA.
#cross-references	MolUD:94187072
#accession	A53988
#status	preliminary
##molecule_type	mRNA
##residues	1-441
##cross-references	GB:120681, NID:g404781, PID:g404782
GENETICS	
#gene	ets-1
CLASSIFICATION	#superfamily transcription factor ets; ets DNA-binding domain homology; ets RII regulatory region homology alternative splicing; DNA binding; nucleus; phosphoprotein; proto-oncogene
KEYWORDS	
FEATURE	
57-130	#domain ets RII regulatory region homology #label ETS2\
97-130	#region helix-loop-helix #status predicted\
337-415	#domain ets DNA-binding domain homology #label ETS\
377-383	#region nuclear location signal
SUMMARY	#length 441 #molecular-weight 50422 #checksum 6268
Query Match	13.2% Score 456; DB 1; Length 441; Best Local Similarity 51.7%; Pred. No. 9,606-65;

	Matches	60;	Conservative	25;	Mismatches	30;	Indels	1;	Gaps	1;
Db	310	DRADLNKDKPVIPAAALAGYTGSGPIQLWQFLLELLTDKSCQSFISWTGDNWFEKLSDPD	369							
Oy	285	ORPOLDPYOILGPTSSRLANPGSGOQLWQFLLELLSDSSNSNCITWECTNGCFKMTDPD	344							
Db	370	EYARRWGKRKKPKMNYEKLRSGLRYYYDKNIHKT-AGKRYVRFVCDLQSLGy	424							
Oy	345	EYARRWGERKSKPMNMYDKLSRALRYYYDKNIMTKVHPPESSMYRYPSDLPYMSsy	400							

Search completed: Wed Apr 7 09:44:31 1999  
Job time : 78 secs.



\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
\*\*\*\*\*  
W E S E R E (TM)  
\*\*\*\*\*

MSearch.p protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed Apr 7 09:44:50 1999; Maspar time 13.89 Seconds  
Molecular output not generated. 923.482 Million cell updates/sec

Title: >US-08-878-177-4  
Description: (1-478) from US08878177.pep  
Perfect Score: 3467  
Sequence: 1 MASTIKELSVSEDSLE.....IYPNRLPAHMPHSLGTY 478

Scoring table: PAM 150  
Gap 11

Searched: 74019 segs, 26840295 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot35  
1:swissprot

Statistics: Mean 49.703; Variance 83.064; scale 0.598

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2608	75.2	462	1	ERG_HUMAN TRANSFORMING PROTEIN E	0.00e+00
2	1776	51.2	452	1	FLI1_HUMAN FLI-1 ONCOGENE (ERGB T	0.00e+00
3	1773	51.1	452	1	FLI1_MOUSE RETROVIRAL INTEGRATION	0.00e+00
4	1740	50.2	453	1	FLI1_XENLA RETROVIRAL INTEGRATION	0.00e+00
5	1571	45.3	272	1	ERG_MOUSE TRANSFORMING PROTEIN E	0.00e+00
6	745	21.5	173	1	ERG_LYTV A ERG PROTEIN HOMOLOG (F	1.36e-144
7	519	15.0	102	1	ETS6_DROME DNA-BINDING PROTEIN D-	1.26e-85
8	459	13.2	64	1	ETS3_DROME DNA-BINDING PROTEIN D-	1.46e-75
9	458	13.2	268	1	ET1B_XENLA C-ETS-1B PROTEIN (XEL-	2.49e-75
10	458	13.2	438	1	ET1A_XENLA C-ETS-1A PROTEIN.	2.49e-75
11	456	13.2	441	1	ETSL_MOUSE C-ETS-1 PROTEIN (P54).	7.28e-75
12	456	13.2	441	1	ETSL_RAT C-ETS-1 PROTEIN (P54).	7.28e-75
13	456	13.2	441	1	ETSL_HUMAN C-ETS-1 PROTEIN (P54).	7.28e-75
14	453	13.1	472	1	ET2A_XENLA C-ETS-2A PROTEIN.	3.63e-74
15	452	13.0	441	1	ETSA_CHICK TRANSFORMING PROTEIN P	6.19e-74
16	452	13.0	454	1	GABA_HUMAN GA BINDING PROTEIN ALP	6.19e-74
17	452	13.0	454	1	GABA_MOUSE GA BINDING PROTEIN ALP	6.19e-74
18	452	13.0	464	1	ELG_DROME DNA-BINDING PROTEIN D-	6.19e-74
19	450	13.0	472	1	ET2B_XENLA C-ETS-2B PROTEIN.	1.81e-73
20	452	13.0	485	1	ETSB_CHICK TRANSFORMING PROTEIN P	6.19e-74
21	447	12.9	479	1	ET2S_CHICK C-ETS-2 PROTEIN.	8.37e-73
22	448	12.9	669	1	MYBE_AVILE P135-GAG-MYB-ETS TRANS	5.26e-73
23	445	12.8	468	1	ETS2_MOUSE C-ETS-2 PROTEIN.	2.61e-72

ALIGNMENTS

RESULT ID	1	STANDARD:	PRT:	462 AA.
AC	ERG_HUMAN P11308;			
DT	01-JUL-1989 (REL. 11, CREATED)			
DT	01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	TRANSFORMING PROTEIN ERG (ERG-2 / ERG-1).			
GN	ERG.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUCHARYOTA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE; 87263429.			
RA	RAO V.N., PAPAS T.S., SHYAM E., REDDY P.;			
RL	SCIENCE 237:635-639(1987).			
RN	[2]			
RP	SEQUENCE OF 100-462 FROM N.A.			
EX	MEDLINE; 87317608.			
RA	REDDY E.S.P., RAO V.N., PAPAS T.S.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 84:6131-6135(1987).			
RN	[3]			
RP	CHROMOSOMAL TRANSLOCATION.			
RX	MEDLINE; 94356859.			
RA	DUNN T., PRAISSMAN L., HAGAG N., VIOLA M.V.;			
RL	CANCER GENET. CITOGENET. 76:19-22(1994).			
RN	[4]			
RP	CHROMOSOMAL TRANSLOCATION.			
EX	MEDLINE; 94243799.			
RA	ICHIKAWA H., SHIMIZU K., HAYASHI Y., OHKI M.;			
RL	CANCER RES. 54:2865-2868(1994).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED			
CC	BY A CHROMOSOMAL TRANSLOCATION T(16;21)(P11;Q22) WHICH INVOLVES			
CC	ERG AND FUS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
DR	EMBL; M17254; G182187; -.			
DR	EMBL; M21535; G182185; -.			
DR	PIR; A29515; TVHDEG. -.			
DR	TRANSFAC; T00265; -.			
DR	TRANSFAC; T00266; -.			
DR	MTM; 165080; -.			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE; PS50061; ETS_DOMAIN_3; 1.			
KW	NUCLEAR PROTEIN; TRANSFORMING PROTEIN; DNA-BINDING; ALTERNATIVE INITIATION; PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION.			



Yy	418	LQPAHNPQKNFVAPHPHPALPVLVTSSFFAANPYNWSPPTGTYENT---	RLPAAHNSHL	474
Dd	449	GSYY	452	
Oy	475	GTYX	478	
RESULT	3	STANDARD:	PRT:	452 AA.
ID	FL11.MOUSE			
AC	P26323;			
DT	01-MAY-1992 (REL. 22, CREATED)			
DT	01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)			
DE	RETROVIRAL INTEGRATION SITE PROTEIN FL1-1.			
GN	FL11 OR FL1-1.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUDAROTA: METAZOA: CHORDATA: VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA: RODENTIA.			
RN	[1]			
RE	SEQUENCE FROM N.A.			
RA	STRAIN-BALB/C; TISSUE--SPLEEN;			
RA	MEDLINE: 91257578.			
RA	BEN-DAVID Y., GIDDENS E.B., LETWIN K., BERNSTEIN A.,			
RA	GENES DEV. 5:908-918(1991).			
RL	[2]			
RP	CHARACTERIZATION			
RX	MEDLINE: 93275657.			
RA	ZHANG L., LEWARCHANDEL V., ROMEO P.-H., BEN-DAVID Y., GREER P.,			
RA	BERNSTEIN A.;			
RL	ONCOGENE 8:1621-1630(1993).			
CC	-1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES			
CC	THE DNA SEQUENCE 5'C[CA]GGAGT-3'.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN BOTH HEMATOPOIETIC AND			
CC	NONHEMATOPOIETIC TISSUES.			
CC	-1- INDUCTION: SPECIFICALLY UP-REGULATED IN LEUKEMIC CLONES WITH			
CC	F-MUZY INSERTIONS UP-STREAM OF THE FL1-1 LOCUS.			
CC	-1- DISEASE: INVOLVED IN ERYTHROLEUKEMIA INDUCTION BY FRIEND MURINE			
CC	LEUKEMIA VIRUS (F-MUZY).			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
DR	EMBL: X59421; G50975; -.			
DR	PIR: S17403; S17403.			
DR	TRANSFAC: T01408; -.			
DR	MGI: 95554; FL1.			
DR	PROSITE: PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE: PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE: PS50061; ETS_DOMAIN_3; 1.			
DR	TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;			
DR	PROTO-ONCOGENE.			
PT	DNA_BIND	281	361	ETS-DOMAIN.
PT	SEQUENCE	452 AA;	51002 MW;	CECCDD4 CRC32;
Query Match		51.1%;	Score 1773;	DB 1; Length 452;
Best Local Similarity		57.2%;	Pred. No. 0.00e+00;	
Matches	277;	Conservative	89;	Mismatches 80; Indels 38; Gaps 24;
Dd	1	MDGTTKALSVSDQSLFDSATYGAHAHLPRADMTAGSGPDYGGPHKINLPPOQEMINQ	60	
Oy	1	MASTIKALSVSEDSOLFECAYGSP-HIAKTEMTASSSSSEYCGTSMSPVPODWLSQ	59	
Dd	61	P-VARNVAREYD--HMNSRSRPDCSVSKNKLYGSGEAMPYNYNMDKKNPPPNM	117	
Oy	60	PPARVTIMECPNPNVNSRNSPDDCSYAKGKNWSSDDNKNMNGYSIMEKH-IIPPNM	118	
Dd	118	TTNERRVIVPADPTLMTQDEHVRWLEMAIKERYGLMEIDTSEFFQNMDDKEJCKNKEDFLR	177	
Oy	119	TTNERRVIVPADPTLMTSDHVRWLEMAIKERYGLPDVILFQINDGKEJCKNKEDFOR	178	
Dd	178	ATSAVNEVLVLSHYLRESSL-LAYTTSHT-DQSRK-NVKE---D---P-S-Y-	222	
Oy	179	LTPSNADILISHLHYLETPPLPHLTSDVDKALQNSPRLMHAHNTGATFFIPNTSV	238	

Db	223	DSV-R--R-G-ANNNNNNSG-LUKS-PLLG-GSG----	TMGNKTEORPPDPDYOIGPT	269
Qy	229	EATQRTTRPDLPLEQARRSAMTSHSPHOTOSKATQPBSSSTVPKTEDEORPDLDPYQILGPT	298	
Db	270	SSRLAMPGSGQIQLMQFLELLELSDSANASCITWGTNGEFTKMTDPDEVARMRGERSKPN	329	
Qy	299	SSRLAMPGSGQIQLMQFLELLELSDSANASCITWGTNGEFTKMTDPDEVARMRGERSKPN	358	
Db	330	MNYDKLSAALYYDYDKNIMTKVHKRKRAIKYFDFHGIQALQPHPTETSMKTP-SDLSYM	368	
Qy	359	MNYDKLSAALYYDYDKNIMTKVHKRPRESS-WKYPSPSDLPYMSYHGRKRAYAKPFHGHQA	417	
Db	369	PSYAHOOKEVFEVSHSSSPMTSSSEFGAASQVCTSPATGATVPNPSPRHPTVHPSHL	448	
Qy	418	LQPAHHPQKMFVAPHPHPALPVTSSSEFFAAPNPVWNSPPTGCTIPTN-T-RLPAHHPSHL	474	
Db	449	GSYX 452		
Qy	475	GTXY 478		
RESULT	4	STANDARD;	PRT;	453 AA.
ID	FL11_XENLA			
AC	PA1157;			
DT	01-FEB-1995 (REL. 31, CREATED)			
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	RETROVIRAL INTEGRATION SITE PROTEIN FL1-1 HOMOLOG.			
GN	FL1.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EUAROTHA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 94206844.			
RA	MEYER D., WOLFE C.M., STEIGLER P., SENAN F., BEFORT N.,			
RA	BEFORT J.J., REMY P.;			
RL	MECH. DEV. 44:109-121(1993).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
DR	EMBL; G65979; G505487; -.			
DR	TRANSFAC; T02067; -.			
DR	PROSITE; PS00345; ETS DOMAIN 1; 1.			
DR	PROSITE; PS00346; ETS DOMAIN 2; 1.			
DR	PROSITE; PS50061; ETS DOMAIN 3; 1.			
KW	TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN.			
FT	DNA_BIND 282 362 ETS-DOMAIN.			
SQ	SEQUENCE 453 AA; 51015 MW; DEB5K90B CRC32;			
Query Match	50.2%;	Score 1740;	DB 1;	Length 453;
Best Local Similarity	56.4%;	Pred. No. 0.00e+00;		
Matches	273;	Conservative 91;	Mismatches 83;	Indels 37; Gaps 25;
Db	1	MDGRTKALSVYSDOSDIFDSAYGASSHLKADMTATGAPGPHKINIPPOODMNG	60	
Qy	1	MASTIKELSVSDEDSQIFECATG-SHLAKTEMTASSSEYIGOTSMSPRYQODMLSQ	59	
Db	61	P-MRVNIRREYE--HMNGRESRPVDCSINCKSLIGGSEGNAMTY-TYMDERKNPPPNM	116	
Qy	60	PPARTIKMECNPNQVNGSRNSPDDCSVAKGKGKMWSSDDVNGYSGYMEKH-IPPNM	118	
Db	117	TNTERVYVPADPALMSQDHHKQMLEAKREYGLVEIDCSLFFONIDCKELCKSKDEFLR	176	
Qy	119	TNTERVYVPADPLMTSDHAKQMLEAVAEYGLVDVILFONIDCKELCKMTKDDFOR	178	
Db	177	STSYINTEVLLSHNTYLRDS--SLIGYNTQANT-DQSSRLT-AKE--D---P--S-Y-	222	
Qy	179	LTPSNADILILSHLYRETPRLPLTSDVDKALQNSPRLMHARNNGATFIPNMSYVP	238	
Db	223	EAV-R--RSG--WGNKSSP-VTKSPPM-G-GTQ-VNKK-S-GDOORSPDPYQILGPT	270	
Qy	239	EATQRTTRPDLPLEQARRSAMTSHSPHOTOSKATQPBSSSTVPKTEDEORPDLDPYQILGPT	298	





Y 285 ORPOLDPQILGPTSSRLANGSGOIQLOMFLLELLSSNSNCITWGTGEEFKMTDPD 344  
Db 367 EVARRMGKRRKPKMANYEKLRSGLRYYYDKNIHKT-AGKRYVRYFVCDLQSLG 421  
Y 345 EVARRMGKRRKPKMANYDKLSRALRYYYDKNIMTKVHPRESSMYKYPSPDLPMSSY 400

RESULT 11  
ID ETS1\_MOUSE STANDARD: PRT: 440 AA.  
AC P27577: 061403;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE C-ETS-1 PROTEIN (P54).  
OS ETS1 OR ETS-1.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
[1]  
SEQUENCE FROM N.A.  
WATSON D.K., SMYTH F.E., SCHWEINFEST C.W., PAPAS T.S.;  
RL (IN) ONCOGENESIS, PAPAS T.S., ED., PP. 221-232, GULF PUBLISHING  
COMPANY, HOUSTON, (1990).  
[2]  
SEQUENCE FROM N.A.  
STRAIN-BALB/C; TISSUE-THYMUS;  
RX MEDLINE: 90299137.  
RA GUNTHER C.V., NYE J.A., BRYNER R.S., GRAVES B.J.;  
RL GENES DEV. 4:667-679(1990).  
[3]  
SEQUENCE FROM N.A.  
STRAIN-BALB/C; TISSUE-THYMUS;  
RX MEDLINE: 90370376.  
RA CHEN J.H.;  
RL ONCOGENE RES. 5:277-285(1990).  
[4]  
STRUCTURE BY NMR OF 332-415.  
RX MEDLINE: 96176767.  
RA DONALDSON L.W., PETERSEN J.M., GRAVES B.J., MCINTOSH L.P.;  
RL EMBO J. 15:125-134(1996).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
-1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
EMBL: M58482; G193192; -.  
DR EMBL: X53953; G50871; -.  
DR EMBL: X55787; G296023; -.  
DR PIR: A30487; A30487.  
DR PDB: 1ETC: 29-JAN-96.  
DR PDB: 1ETD: 28-JAN-96.  
DR TRANSFAC: T0011; -.  
DR MGD: MGI:95455; ETS1.  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
KM PHOTO-ONCOGENE; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING;  
KW PHOSPHORYLATION; 3D-STRUCTURE.  
FT DNA\_BIND 335 415  
FT CONFLICT 28 28 D->E (IN REF. 1 AND 3).  
FT CONFLICT 37 37 L->S (IN REF. 3).  
FT CONFLICT 51 52 AT->SY (IN REF. 3).  
FT CONFLICT 55 55 G->P (IN REF. 1).  
FT CONFLICT 55 55 G->P (IN REF. 1).  
FT CONFLICT 63 63 L->R (IN REF. 3).  
FT CONFLICT 74 74 E->D (IN REF. 3).  
FT CONFLICT 96 96 Q->H (IN REF. 3).  
FT CONFLICT 105 105 L->V (IN REF. 3).  
FT CONFLICT 157 157 D->V (IN REF. 3).  
FT CONFLICT 211 211 Q->R (IN REF. 3).  
FT CONFLICT 217 217 Q->E (IN REF. 3).  
FT CONFLICT 225 225 A->R (IN REF. 3).  
FT CONFLICT 234 234 D->N (IN REF. 3).  
FT CONFLICT 360 360 G->C (IN REF. 3).  
FT CONFLICT 383 383 K->S (IN REF. 3).  
FT CONFLICT 392 392 G->A (IN REF. 3).

FT CONFLICT 408 409 KR->NA (IN REF. 3).  
FT CONFLICT 413 413 R->A (IN REF. 3).  
SQ SEQUENCE 440 AA; 50201 MW; 18E90F01 CRC32;  
Query Match 13.2%; Score 456; DB 1; Length 440;  
Best Local Similarity 51.7%; Pred. No. 7.28e-75;  
Matches 60; Conservative 25; Mismatches 30; Indels 1; Gaps 1;  
Db 310 DRADLNKDKPYIPAAALAGYTGSGPIQLOMFLLELLDSCQSFISWTGDEWFKLSDDP 369  
Y 285 ORPOLDPQILGPTSSRLANGSGOIQLOMFLLELLSSNSNCITWGTGEEFKMTDPD 344  
Db 370 EVARRMGKRRKPKMANYEKLRSGLRYYYDKNIHKT-AGKRYVRYFVCDLQSLG 424  
Y 345 EVARRMGKRRKPKMANYDKLSRALRYYYDKNIMTKVHPRESSMYKYPSPDLPMSSY 400

RESULT 12  
ID ETS1\_RAT STANDARD: PRT: 441 AA.  
AC P41156;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE C-ETS-1 PROTEIN (P54).  
OS ETS1 OR ETS-1.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE: 94187072.  
RA BELLA-COSA A., DATTA K., BEAR S.E., PATRIOTIS C., LAZO P.A.,  
RA COPELAND N.G., JENKINS N.A., TSICHLIS P.N.;  
RL J. VIROL. 68:2320-2330(1994).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
-1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
EMBL: L20681; G404782; -.  
DR PIR: A53988; A53988.  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
KM PHOTO-ONCOGENE; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING;  
KW PHOSPHORYLATION.  
FT DNA\_BIND 335 415  
FT SEQUENCE 441 AA; 50423 MW; 1C04335A CRC32;  
Query Match 13.2%; Score 456; DB 1; Length 441;  
Best Local Similarity 51.7%; Pred. No. 7.28e-75;  
Matches 60; Conservative 25; Mismatches 30; Indels 1; Gaps 1;  
Db 310 DRADLNKDKPYIPAAALAGYTGSGPIQLOMFLLELLDSCQSFISWTGDEWFKLSDDP 369  
Y 285 ORPOLDPQILGPTSSRLANGSGOIQLOMFLLELLSSNSNCITWGTGEEFKMTDPD 344  
Db 370 EVARRMGKRRKPKMANYEKLRSGLRYYYDKNIHKT-AGKRYVRYFVCDLQSLG 424  
Y 345 EVARRMGKRRKPKMANYDKLSRALRYYYDKNIMTKVHPRESSMYKYPSPDLPMSSY 400

RESULT 13  
ID ETS1\_HUMAN STANDARD: PRT: 441 AA.  
AC P14921;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE C-ETS-1 PROTEIN (P54) (C-ETS-1A AND C-ETS-1B).  
OS ETS1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE: 89083219.





CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

DR EMBL; M22462; G211753; -

DR	EMBL; X13026; G63180; -
DR	EMBL; X13027; G63383; -

DR ENBU, A31285: TYCHTE.

DR PIR; S07625; S07625.

DR      PIR; S29132; S29132.

DR TRANSFAC; T00114; -.

DR	PROSITE; PS00345; ETS_DOM
DR	PROSITE; PS00346; ETS_DOM

```

DR      PROSTIE; ETS-DOM
DB      PS50061; ETS-DOM

```

PROTO-ONCOGENE; DNA-BINDING

FT	DNA_BIND	335	415
----	----------	-----	-----

FT	CONFLICT	117	117
----	----------	-----	-----

SQ SEQUENCE 441 AA; 50326

Query Match	13 08
1	13 08
2	13 08
3	13 08
4	13 08
5	13 08
6	13 08
7	13 08
8	13 08
9	13 08
10	13 08
11	13 08
12	13 08
13	13 08
14	13 08
15	13 08
16	13 08
17	13 08
18	13 08
19	13 08
20	13 08
21	13 08
22	13 08
23	13 08
24	13 08
25	13 08
26	13 08
27	13 08
28	13 08
29	13 08
30	13 08
31	13 08
32	13 08
33	13 08
34	13 08
35	13 08
36	13 08
37	13 08
38	13 08
39	13 08
40	13 08
41	13 08
42	13 08
43	13 08
44	13 08
45	13 08
46	13 08
47	13 08
48	13 08
49	13 08
50	13 08
51	13 08
52	13 08
53	13 08
54	13 08
55	13 08
56	13 08
57	13 08
58	13 08
59	13 08
60	13 08
61	13 08
62	13 08
63	13 08
64	13 08
65	13 08
66	13 08
67	13 08
68	13 08
69	13 08
70	13 08
71	13 08
72	13 08
73	13 08
74	13 08
75	13 08
76	13 08
77	13 08
78	13 08
79	13 08
80	13 08
81	13 08
82	13 08
83	13 08
84	13 08
85	13 08
86	13 08
87	13 08
88	13 08
89	13 08
90	13 08
91	13 08
92	13 08
93	13 08
94	13 08
95	13 08
96	13 08
97	13 08
98	13 08
99	13 08
100	13 08

Query	Local	Similarity
Best	50.98	

Matches 59; Conservative

310 DRADMNKDKPVIIPAALAGY

285 OPDOPVOTTEPTSCPIAN

200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600 610 620 630 640 650 660 670 680 690 700 710 720 730 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560 4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600 5610 5620 5630 5640 5650 5660 5670 5680 5690 5700 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300 6310 6320 6330 6340 6350 6360 6370 6380 6390 6400 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600 6610 6620 6630 6640 6650 6660 6670 6680 6690 6700 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000 7010 7020 7030 7040 7050 7060 7070 7080 7090 7100 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200 7210 7220 7230 7240 7250 7260 7270 7280 7290 7300 7310 7320 7330 7340 7350 7360 7370 7380 7390 7400 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500 7510 7520 7530 7540 7550 7560 7570 7580 7590 7600 7610 7620 7630 7640 7650 7660 7670 7680 7690 7700 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800 7810 7820 7830 7840 7850 7860 7870 7880 7890 7900 7910 7920 7930 7940 7950 7960 7970 7980 7990 8000 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100 8110 8120 8130 8140 8150 8160 8170 8180 8190 8200 8210 8220 8230 8240 8250 8260 8270 8280 8290 8300 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400 8410 8420 8430 8440 8450 8460 8470 8480 8490 8500 8510 8520 8530 8540

Db 370 EVARRWGKRNKPKMNYEKI

Search completed: Wed Apr 7 09:46:02 1999  
Job time : 72 secs.



Qy	181	PSYNADILSLHRLRE-----R-----GATFIFPNPNVYPEA	213
Db	181	PSYNADILSLHRLREPHILTSDDVDKQASPLMHARNGATFIFPNPNVYPEA	240
Qy	181	PSYNADILSLHRLRE-----R-----GATFIFPNPNVYPEA	213
Db	241	TQRTITTPDLEEDARRSAMTSHSHPTQSKATOPSSSVKTEEDQRPQDLPYQILGPTSS	3000
Qy	214	TQRTITTPDLEEDARRSAMTSHSHPTQSKATOPSSSVKTEEDQRPQDLPYQILGPTSS	273
Db	301	RLANPGSQIOLMOLFLELLSDSSNSNCITWEGNGEFKMTDPDEVARRNGERKSKNMN	366
Qy	274	RLANPGSQIOLMOLFLELLSDSSNSNCITWEGNGEFKMTDPDEVARRNGERKSKNMN	333
Db	361	YDKLSRALRYYYDXNIMTKVHGKRAYKEDFEGIAQALQHPPESSMYKTPSDLPYWSY	420
Qy	334	YDKLSRALRYYYDXNIMTKVHGKRAYKEDFEGIAQALQHPPESSMYKTPSDLPYWSY	393
Db	421	HAHPQKNMFVAPHPALPVTSSSFEAAPNPWNSPTGIGYNTBCLPAHNPSHLGITY	478
Qy	394	HAHPQKNMFVAPHPALPVTSSSFEAAPNPWNSPTGIGYNTBCLPAHNPSHLGITY	451

```

FEATURE
1-231,256-486
100-231,256-486
126-200

#domain ets RII regulatory region homology #label ETS2\

ACCESSIONS
A94294; A94178; I58410; A28041; A23515

REFERENCE
#authors
#journal
#title
#cross-references MUID:87263429

#molecule_type mRNA
#accession A94294
##residues 1-231,256-486 ##label REA

REFERENCE
#cross-references GB:M17254; VIND:g182186; PID:g182187
A94178
#authors
#journal
#title
#cross-references MUID:87317608

#molecule_type mRNA
#residues 100-231,256-486 ##label REB

REFERENCE
#cross-references GB:M21535; NID:g182182; PID:g182185
I58410
#authors
#journal
#title
#cross-references MUID:94119611

#molecule_type mRNA
#accession I58410
##residues 230-259 ##label RES

#status
#molecule_type mRNA
##residues
#cross-references GB:S68130; NID:g544960; PID:g544961

#gene
GDB:ERC
#cross-references GDB:119884; OMIM:165080
#map_position 21q22.2-21q22.2

CLASSIFICATION
#paraphrasing transcription factor erg; ets DNA-binding domain
homology; ets RII regulatory region homology
alternative splicing; DNA binding; nucleus; proto-oncogene;
transcription factor; transforming protein

KEYWORDS
#product transforming protein erg-2 #status predicted
#label EG2\
#product transforming protein erg-1 #status predicted
#label EG1\

#domain ets RII regulatory region homology #label ETS2\

```

```

320-398      #domain ets DNA-binding domain homology #label ETS
SUMMARY      #length 486      #molecular-weight 54608      #checksum 9577

Query Match      87.9%      Score 2887; DB 1;      Length 486;
Best Local Similarity 89.3%;      Pred. No. 0.00e+00;
Matches 427;      Conservative 20;      Mismatches 3;      Indels 28;      Gaps 3

```

	9	AALKKALTVSDDSLFECVAGTPIPLATEMTASTSSSGYGTSMKSRVRPOOMLSDPP	68
	10	:::     :::     :::     :::     :::     :::     :::	
Oy	2	ASTIKELALVSVEDQSLFECAYGSPILAKTEMTASTSSSEFYGTSMKSRVRPOOMLSDPP	61
Db	69	ARVTIKMECPNPGVNSRNSPDECSYAKGKKMYGPDIVGNMYSYMEEKHMPPNMTTN	128
	70	:::     :::     :::     :::     :::     :::     :::	
Oy	62	ARVTIKMECPNPGVNSRNSPDCSYAKGKKMYGPDIVGNMYSYMEEKHMPPNMTTN	121
Db	129	ERRIYVADPTLSTGTHVQMILEMANKEKGEDNVNILLFPONTDCKELCKMKRDFORLTP	188
	130	:::     :::     :::     :::     :::     :::     :::	
Oy	122	ERRIYVADPTLSTGTHVQMILEMAWKEGGLDPVDILLFPONTDCKELCKMKRDFORLTP	181
Db	189	SYNADILLSHLARTPTPHLLTSDVDVALONSDPRLMHARNTGGAEIFPNTSVYPEAT	248
	190	:::     :::     :::     :::     :::     :::     :::	
Oy	182	SYNADILLSHLARTPE-----R---CATIFPFNTSVYPEAT	214
Db	249	QRITTRPDLEYPERRASANTGCHGPPOSKAAPSPSTVPKTEDORPOLDYQLGPTSS	308
	250	:::     :::     :::     :::     :::     :::     :::	
Oy	215	QRITTRPDLEYPARRASANTSHHP--QSKAAPSPSTVPKTEDORPOLDYQLGPTSS	273
Db	309	RLANPGGGQIQLOQFLLELLSDSNSNCITTWGCTNGEERKMTPDDEVARMERKSRRPMN	368
	310	:::     :::     :::     :::     :::     :::     :::	
Oy	274	RLANPGGGQIQLOQFLLELLSDSNSNCITTWGCTNGEERKMTPDDEVARMERKSRRPMN	333
Db	369	YDKLSRALRYYYDNKIMTKVHGKRAYKEDFHGIQAOLAPHPRESSLYKYPSDLPLYGGSY	428
	370	:::     :::     :::     :::     :::     :::     :::	
Oy	334	YDKLSRALRYYYDNKIMTKVHGKRAYKEDFHGIQAOLAPHPRESSMYKYPSDLPYMSY	393
Db	429	HAHQKNFAFAPRPALPYTSSSFPAAPRVNNSPTGGIGYPTRTLRPTSHMSHLGTYX	486
	430	:::     :::     :::     :::     :::     :::     :::	
Oy	394	HAHQKNFAFAPRPALPYTSSSFPAAPRVNNSPTGGIGYPTRTLRPAAHMPSHLGTYX	451
RESULT	3	137565 #type complete	
ENTRY		transforming protein fl1, long splice form - human	
TITLE		Friend leukemia integration protein 1; transcription factor	
ALTERNATE_NAMES		ERGB	
ORGANISM		#formal_name Homo sapiens #common_name man	
DATE		04-Oct-1996 #sequence_revision 30-May-1997 #text_change	
ACCESSIONS		137565: S29843: S35506: A49000: A49015: IS4170	
REFERENCE		I37378	
#authors		Delattre, P.; Zucman, J.; Plougastel, B.; Desmaze, C.; Molot,	
		T.; Peter, M.; Kovar, H.; Jobert, I.; de Jong, P.;	
		Rouleau, G.; Aurias, A.; Thomas, G.	
#journal		Nature (1992) 359:163-165	
#title		Gene fusion with an ETS DNA-binding domain caused by	
		chromosome translocation in human tumours.	
#cross-references		NCBI:92396239	
#accession		I37565	
#status		preliminary: translated from GB/EMBL/DBJ	
##molecule_type		mRNA	
##residues		1-452 ##label RES	
##cross-references		EMBL:X67001; NID:g32529; PID:g32530	
REFERENCE		S29843	
#authors		Homas, R.; May, W.; Denny, C.; Roskind, W.; Moore, J.; Makl,	
		R.A.; Beck, E.; Klemz, M.J.	
#journal		Biochim. Biophys. Acta (1993) 1172:155-158	
#title		Human FLI-1 localizes to chromosome 11q24 and has an aberrant	
		transcript in neuroepithelioma.	
#cross-references		NCBI:93176799	
#accession		S29843	
##molecule_type		mRNA	
##residues		1-66, '/', '70-129, 'A', '131-132, 'V', '134-322, 'O', '324-425,	
		427-452 ##label HNOI	

```

#cross-references EMBL:M93255
#note
    The authors translated the codon GTG for residue 69 as
    glu

REFERENCE
#authors
    Hromas, R.A.; May, M.; Denny, C.; Reskind, W.; Moore, J.;
    Maki, R.A.; Beck, E.; Klemz, M.J.
#submissions
    submitted to the EMBL Data Library, May 1993
#description
    Human Fli-1, an ETS oncogene family member preferentially
    expressed in hematopoietic cells, is a sequence-specific DNA
    binding protein.

#accession
    S35506
#molecule_type
    mRNA
#residues
    1-68, 'V', 70-76, 78-129, 'A', 131-132, 'V', 134-322, 'Q',
    324-425, 427-452 ##label HRO2
#cross-references EMBL:M93255; NID:g182659; PID:g182660
A49000
#authors
    Prasad, D.D.; Rao, V.N.; Reddy, E.S.
#journal
    Cancer Res. (1992) 52:5833-5837
#title
    Structure and expression of human Fli-1 gene.
#cross-references MUID:93007976
#accession
    A49000
#status
    Preliminary
#molecule_type
    mRNA
#residues
    1-322, 'Q', 324-452 ##label PRA
#cross-references GB:S45205; NID:g257353; PID:g257354
#note
    Sequence extracted from NCBI backbone (NCBIN:115336,
    NCBI:P:115337)
A49015
#authors
    Watson, D.K.; Smyth, F.E.; Thompson, D.M.; Cheng, J.Q.;
    Testa, J.R.; Papas, T.S.; Seth, A.
#journal
    Cell Growth Differ. (1992) 3:705-713
#title
    The ERGB/Fli-1 gene: isolation and characterization of a new
    member of the family of human ETS transcription factors.
#cross-references MUID:93075640
#accession
    A49015
#status
    Preliminary
#molecule_type
    mRNA
#residues
    1-390, 'Q', 392-425, 427-452 ##label WAT
#cross-references GB:M98833; NID:g182168; PID:g182169
#note
    Sequence extracted from NCBI backbone (NCBIN:119390,
    NCBI:P:119392)
I54170
#authors
    Baud, V.; Lipinski, M.; Rassart, E.; Polluin, L.; Bergeron,
    D.
#journal
    Genomics (1991) 11:223-224
#title
    The human homolog of the mouse common viral integration
    region, Fli1, maps to 11q23-q24.
#cross-references MUID:92112219
#accession
    I54170
#status
    preliminary; nucleic acid sequence not shown; translated
    from GB/EMBL/DBJ
##molecule_type
    DNA
#residues
    1-6 ##label RZ2
#cross-references GB:L47616; NID:g1000863; PID:g1000864
GDB:FLI1
#gene
    GDB:FLI1
#cross-references GDB:127565; OMIM:193067
#map_position
    11q24.1-11q24.3
#introns
    77/3
#superfamily
    superfamily transcription factor erg; ets DNA-binding domain
#homology
    homology: ets RII regulatory region homology
#alternative_splicing
    alternative splicing: DNA binding; nucleus; proto-oncogene;
    transcription factor; transforming protein
FEATURE
118-192
283-361
SUMMARY
Query Match 71.5%; Score 23.48; DB 1; Length 452;
Best local Similarity 67.0%; Pred. 0.00e+00;
Matches 306; Conservativity 85; Mismatches 55; Indels 11; Gaps 8;
1 MDGTRKELSVSDOSLFDSAYGAANAHLPKADMTASGSPDYQCPKRIINLPPOEQWING 60

```

QY	1	MASTIKREALSVSEDO	LECAI	GP	-HIAKTEMTASSSESEYQ	TSMKSPRV	QO	DWLSQ	59
QY	61	P-VRVNVKREYD	-	-HNGSRES	PVDCSVKSC	KLVCAGSCSPNNYNS	SYMDERK	NGPP	117
QY	60	PPARTIKRECP	NPQVNGSR	NSPDCD	SAKGGKMG	SSSDNVGNN	GSYMEKH	-IPP	118
QY	118	TTRNRRVVPAD	PTLMTD	OEHVQW	LEMAIK	ESLMEIDTS	FEQNM	DKEK	177
QY	119	TTRNRRVVPAD	PTLMTSD	HVRQW	LEMAVKE	GLPDVDIL	LFQND	KELCK	178
QY	178	ATTLYNTEV	ELVLSHL	RE-SULAY	-NTTSHTD	QSSRLSVK	KDP	SYDSVR	235
QY	179	LPSPYNADIL	LSHLH	LRERGA	TFFIPNTSV	YPEAQRIT	TRDE	LEYEQAR	238
QY	236	SGLNSPPLG	QAOTISK	NTKEOR	POPP	POYOLIG	TSSRLN	PNP	295
QY	239	PQOKATPO	SSS-TYP	KTDQR	POLD	POYLIG	TSSRLN	PNP	297
QY	296	NASCITW	ETNGE	FKMTD	DEVAR	BMGEKRS	KPNMYDKL	SRLARY	355
QY	298	NSNCITW	ETNGE	FKMTD	DEVAR	BMGEKRS	KPNMYDKL	SRLARY	357
QY	356	YAYKED	FGHIA	LOLPH	PRESS	KYKPS	SDISY	WPSY	415
QY	358	YAYKED	FGHIA	LOLPH	PRESS	KYKPS	SDLPY	WPSY	417
QY	416	FGAASQ	WTSPT	GTGIP	NPVNR	PHRTPH	PSH	GSXY	452
QY	418	FAAPNP	WNSPT	GTGIP	NPVNR	PHRTPH	PSH	GSXY	451
RESULT ENTRY	4	S17403	#type complete						
TITLE			transforming protein fl1 - mouse						
ALTERNATE_NAMES			Friend leukemia integration protein 1; transcription factor ERB						
ORGANISM			formal_name Mus musculus #common_name house mouse						
DATE			31-Dec-1993 #sequence_rev10 31-Dec-1993 #text_change 05-Sep-1997						
ACCESSIONS			S17403						
REFERENCE			S17403						
#authors			Ben-David, Y.; Giddens, E.B.; Letwin, K.; Bernstein, A.						
#journal			Genes Dev. (1991) 5:908-918						
#title			Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, fl1-1, closely linked to c-ets-1.						
#cross-references			accession S17403						
GENETICS			#residues 1-452 #label BEN						
CLASSIFICATION			#cross-references GB:X59421; NID:g50974; PID:g50975						
KEYWORDS			fl1 #superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology						
FEATURE			DNA binding; nucleus; proto-oncogene; transcription factor; transforming protein						
SUMMARY			#domain ets RII regulatory region homology #label ETS2\ #domain ets DNA-binding domain homology #label ETS						
Query Match			Best Local Similarity 71.2%; Score 2339; DB 1; Length 452;						
Matches			305; Conservative 85; Mismatches 56; Indels 11; Gaps 9;						
DB	1	MDGITK	REALSVSD	OSLSD	SAVGA	AHHPRKADMT	ASGSPD	YGCPH	KNLPP
QY	1	MASTIK	REALSVSE	DO	LECAI	GP	-HIAKTEMT	ASSSEY	Q
DB	61	P-VRVNVKREYD	-	-HNGSRES	PVDCSVKSC	KLVCAGSCSPNNYNS	SYMDERK	NGPP	117
QY	60	PPARTIKRECP	NPQVNGSR	NSPDCD	SAKGGKMG	SSSDNVGNN	GSYMEKH	-IPP	118
DB	118	TTRNRRVVPAD	PTLMTD	OEHVQW	LEMAIK	ESLMEIDTS	FEQNM	DKEK	177
QY	119	TTRNRRVVPAD	PTLMTSD	HVRQW	LEMAVKE	GLPDVDIL	LFQND	KELCK	178
DB	178	ATTLYNTEV	ELVLSHL	RE-SULAY	-NTTSHTD	QSSRLSVK	KDP	SYDSVR	235
QY	179	LPSPYNADIL	LSHLH	LRERGA	TFFIPNTSV	YPEAQRIT	TRDE	LEYEQAR	238
DB	236	SGLNSPPLG	QAOTISK	NTKEOR	POPP	POYOLIG	TSSRLN	PNP	295
QY	239	PQOKATPO	SSS-TYP	KTDQR	POLD	POYLIG	TSSRLN	PNP	297
DB	296	NASCITW	ETNGE	FKMTD	DEVAR	BMGEKRS	KPNMYDKL	SRLARY	355
QY	298	NSNCITW	ETNGE	FKMTD	DEVAR	BMGEKRS	KPNMYDKL	SRLARY	357
DB	356	YAYKED	FGHIA	LOLPH	PRESS	KYKPS	SDISY	WPSY	415
QY	358	YAYKED	FGHIA	LOLPH	PRESS	KYKPS	SDLPY	WPSY	417
DB	416	FGAASQ	WTSPT	GTGIP	NPVNR	PHRTPH	PSH	GSXY	452
QY	418	FAAPNP	WNSPT	GTGIP	NPVNR	PHRTPH	PSH	GSXY	451

QY	60	PARATYIKMECNQVNGSNPSDDCSVAAGKMGVSSSDVNGMYSYMEKH-1	IPPPNM	118
Db	118	TNERRVIVPADETLWTOEHVROWLEMAIKGYLMEITDSFFQMDGKEICKNKEDFLR	177	
QY	119	TNERRVIVPADPTLSTBHVROWLEMAIKGYLMEITDSFFQMDGKEICKNKEDFLR	178	
Db	178	ATSAVTEVLLSHLSLRE-SSLLAY-NTSHTDQSSRLVWKEDEPSYDSVRGGMNNMN	235	
QY	179	LPPSYADILLSHLHRELGATFIEPNTSYVEATQRTTRPDLPEQARRSAMTSHS	238	
Db	236	SGLNKSPLLGSGOTGKNTKRPORPOPPYOLIGFTSSRLNPGSGOILMOFILELLSDA	295	
QY	239	PROSKATOPSSS-TVKTEDEQRQPOLDYOLIGFTSSRLNPGSGOILMOFILELLSDS	297	
Db	296	NASCITWETNGEFTKMTDPEVARWGERKSKENPMNDKLSRALRYXXDKNMTKVGRK	355	
QY	298	NSNCITWETNGEFTKMTDPEVARWGERKSKENPMNDKLSRALRYXXDKNMTKVGRK	357	
Db	356	YAKDFPHGIAQLOPHPIETSNKYKIPSDISYMPSTYAHQOKNVPVSHPSMPTVSSSF	415	
QY	358	YAKDFPHGIAQLOPHPIETSNKYKIPSDISYMPSTYAHQOKNVPVSHPSMPTVSSSF	417	
Db	416	FGASQOTWTSPTAGIYPNPSPVPHHPNTHVPSHGSYX	452	
QY	418	FAPNPYMSPTGIGIYPN-T--RLPAAHMPSHGITY	451	
RESULT	5	S49013	#type complete	
ENTRY		transforming protein fl1 - African clawed frog		
TITLE		Friend leukemia integration protein 1; transcription factor		
ALTERNATE_NAMES		ERGB		
ORGANISM		form1_name Xenopus laevis #common_name African clawed frog		
DATE		07-May-1995 #sequence_revision 23-May-1997 #text_change		
ACCESSIONS		S49013		
REFERENCE		S49013		
#authors		Moyor, D.; Wolff, C.M.; Stogler, P.; Senan, F.; Befort, N.;		
#journal		Befort, J.J.; Remy, P.		
#file		Mech. Dev. (1993) 44:109-121		
		X1-file, the Xenopus homolog of the fl1-1 gene, is expressed		
		during embryogenesis in a restricted pattern evocative of		
		neural crest cell distribution.		
#accession		S49013		
#status		preliminary		
##molecule_type		mRNA		
##residues		1-453 #label MEY		
##cross-references		EMBL:X65979; NID:g505486; PID:g505487		
NETICS				
#gene		fl1		
CLASSIFICATION		superfamily transcription factor erg; ets DNA-binding domain		
KEYWORDS		homology; ets RII regulatory region homology		
		DNA binding; nucleus; proto-oncogene; transcription factor;		
		transforming protein		
FEATURE				
117-191		#domain ets RII regulatory region homology #label ERS2\		
284-362		#domain ets DNA-binding domain homology #label ERS		
SUMMARY		#length 453 #molecular-weight 51015 #checksum 774		
Query Match		70.18; Score 2305; DB 1; Length 453;		
Best Local Similarity		65.98; Pred. No. 0.00e+00;		
Matches		302; Conservative 85; Mismatches 59; Indels 12; Gaps 9;		
Db	1	MDGTKEALSVDDSLDPSAYAGASHLSKADMTAANDPYGGPHKINIPPOODMNG	60	
QY	1	MASTIEALSVSDEDSLEFCATG-SHLAKTEKTAASSSEYGGTSMSPVPPOODMSQ	59	
Db	61	P-MRVVIRKEEY--HNHGRSESPVDSINCKSLIGSESENNAMY-TYDEKNPPPPNM	116	
QY	60	PARATYIKMECNQVNGSNPSDDCSVAAGKMGVSSSDVNGMYSYMEKH-1	118	
Db	117	TNERRVIVPADPALSODHVROWLEMAIKGYLMEITDSFLPONDKEICKNKEDFLR	176	

Qy	119	TTNERRIYVADPTLWSTDHVRQMLEMAKKEVGLPVDILLFQNIIDGKELCKMTDDPQR	178
Db	177	STSYNTNEVLISHLNYLRDSSSLGY-NTQAHTDOSSRLTAKEDPSYEAARRSGNSMS	235
Qy	179	LTPSYNADILLSHLYRERGAATFIPNTSVYPEAQRITTPRDLPEQARRSAM-TSHS	237
Db	236	SPVTKSPRMGSTQVNVKNSGDQKSSQDDPVOILGPTSSRLANGSOIQULMQLLELLSDS	295
Qy	238	HPQTSKATQPESSSTVPKT-EDORPQDLPVOILGPTSSRLANGSOIQULMQLLELLSDS	296
Db	296	SNASCIITWEGNGEKKTDDEVEAARRGERKSPNNYNYKLSRALRYDYKDSIMTKVKGK	355
Qy	297	SNSCITITWEGNGEKKTDDEVEAARRGERKSPNNYNYKLSRALRYDYKDSIMTKVKGK	356
Db	356	RYAKKDFPHGIAQLQPHPTDSMTXYXSEFSYMPSPYSHQOKVNFVPSHPSSPVYTSYG	415
Qy	357	RYAKKDFPHGIAQLQPHPTDSMTXYXSEFSYMPSPYSHQOKVNFVPSHPSSPVYTSYG	416
Db	416	FFGATSPYVNSPSANITPVPNPVPHRPTVHOSHLGCFY	453
Qy	417	FFAAPNPYVNSPTGIGYPT--RLPAHMPHSLGTYX	451
RESULT	6		
ENTRY		S29844	#type complete
TITLE			transforming protein fl1, short splice form - human
ALTERNATE_NAMES			friend leukemia integration protein 1; transcription factor ERG
ORGANISM			#formal name Homo sapiens #common name man
DATE			02-Dec-1993 #sequence_revision 23-May-1997 #ctx_change 02-Sep-1997
ACCESSIONS		S29844	
REFERENCE		S29843	
#authors		Hromas, R.; May, W.; Denny, C.; Raabkind, W.; Moore, J.; Makl, R.A.; Beck, E.; Klemenz, M.J.	
#journal		Biochim. Biophys. Acta (1993) 1172:155-158	
#file		Human FL1-1 localizes to chromosome 11Q24 and has an aberrant transcript in neuroepithelioma.	
#cross-references		MUID:9316799	
#accession		S29844	
#status			preliminary; nucleic acid sequence not shown; translation not shown
GENETICS			
#gene		GDB:FL11	
#molecule_type		mRNA	
#residues		1-385	#label HRO
#cross-references		EMBL:M93255; NID:918265; PID:918261	
#note			the nucleotide sequence was submitted to the EMBL Data Library, May 1992
CLASSIFICATION			
#map_position		11q24.1-11q24.3	
KEYWORDS			homology; ets RII regulatory region homology alternative splicing; DNA binding; nucleus; proto-oncogene; transcription factor; transforming protein
FEATURE			
52-126			#domain ets RII regulatory region homology #label ETS2\
217-295			#domain ets DNA-binding domain homology #label ETS
SUMMARY			
Query Match		60.0%; Score 1973; DB 1; Length 385;	
Best Local Similarity		67.8%; Pred. No. 0.00e+00;	
Matches	257; Conservative	68; Mismatches 46; Indels 8; Gaps 6;	
Db	10	ARSDVDCSVSKCKLVGSGESNPMTNINSMDEKNCPPPPNNTTNERRIYVADATLTQTQ	69
Qy	78	SRNSDDCQSAVAKGKMYSSDDNMGNGSYMEKH-IPPNMTTNERRIYVADPTLWST	136
Db	70	EHVQWQLEMAIKKESLWEIDTSPFQNDGKEICKNKEDFLATLTLYNEVLISHLSTLR	129
Qy	137	DHVRQMLEMAKKEVGLPVDILLFQNIIDGKELCKMTDDPQRITTPSYNADILLSHLYLR	196
Db	130	E-SSILAV-NTTSTHDDSSRLSYKEDPSYSVYSGRAGMGNNNSGLNKSPPLGAGDTISKN	187

QY	197	ENGATFIFPNTSVPEATQRTITRPDLPEEQARNSAMTSHSTQSKATQPSSS-ITPKT	255
QY	198	TEORPOPDYQILGPTSSRLANPGSGQIQLOMQLLELLSDSANASCTTGECTNGEEMTD	247
QY	256	EQORPOLDPYQILGPTSSRLANPGSGQIQLOMQLLELLSDSSNSNCITTEGTNGEEMTD	315
DB	248	PDEVARRMGORSKPMNNDKLSRALRYDYDKNIMTKVHKGRAYKFDHFGIAQALQHP	307
QY	316	PDEVARRMGERSKPMNNDKLSRALRYDYDKNIMTKVHKGRAYKFDHFGIAQALQHP	375
DB	308	TESSMYKYSDSISTMYSYSAHQKVFPVPPSPSMPTSSFFGAASQYTS-TGCIYPN	366
QY	376	PESSMYKYSDLPYMSSTYAHAPQKNFVAHPALPYTSSFFFAAPRPYNSTPGGIYPN	435
DB	367	PVNPHPNTHVPSHLGSIY 385	
QY	436	T---RLPAHMPSHLCTYY 451	
RESULT	7		
TITLE		A54617 #type fragment	
ALTERNATE_NAMES		transcription factor erg - mouse (fragment)	
ORGANISM		Ig heavy chain enhancer-binding protein	
DATE		#normal_name Mus musculus #common_name house mouse	
		25-Oct-1994 #sequence_revision 18-Nov-1994 #text_change	
		30-May-1997	
ACCESSIONS		A54617	
REFERENCE		A54617	
#authors		Rivera, R.R.; Stulver, M.H.; Steenbergen, R.; Murre, C.	
#journal		Mol. Cell. Biol. (1993) 13:7163-7169	
#title		Ets proteins: new factors that regulate immunoglobulin heavy-chain gene expression.	
#cross-references		MUID:94019387	
#accession		A54617	
#status		preliminary	
#molecule_type		mRNA	
#residues		1-272 #label R1Y	
#experimental_source		pre-B-cell line 22D6	
#note		sequence extracted from NCBI backbone (NCBIN:138523, NCBI:138524)	
CLASSIFICATION		#superfamily transcription factor erg; ets DNA-binding domain homology	
KEYWORDS		homoology; ets R1I regulatory region homology	
		DNA binding; phosphoprotein; proto-oncogene; transcription factor; transforming protein	
FEATURE			
140-218			
SUMMARY		#domain ets DNA-binding domain homology #label ETS	
		#length 272 #checksum 3754	
Query Match		46.7%; Score 1535; DB 2; Length 272;	
Best Local Similarity		94.5%; Pred. No. 8,79e-296;	
Matches 208; Conservative 9; Mismatches 2; Indels 1; Gaps 1;			
DB	53	GAATFPNVSYPDEAQRITTRPDLYEPPRRSAGMCHSLTQSKRAQSPSAVPTED	112
QY	199	GAATFPNVSYPDEAQRITTRPDLYEPPRRSAGMCHSLTQSKRAQSPSAVPTED	257
DB	113	QRPOLDPYQILGPTSSRLANPGSGQIQLOMQLLELLSDSSNSNCITTEGTNGEEMTD	172
QY	256	QRPOLDPYQILGPTSSRLANPGSGQIQLOMQLLELLSDSSNSNCITTEGTNGEEMTD	317
DB	173	EVARRMGERSKPMNNDKLSRALRYDYDKNIMTKVHKGRAYKFDHFGIAQALQHP	232
QY	318	EVARRMGERSKPMNNDKLSRALRYDYDKNIMTKVHKGRAYKFDHFGIAQALQHP	377
DB	233	SSLKYRPSDLPYMGSHAHPOKKNFSPHPALPYTSSSF 272	
QY	378	SSMYRPSDLPYMSSTYAHAPQKNFVAHPALPYTSSSF 417	
RESULT	8		
ENTRY		A56646 #type fragment	
TITLE		transcription factor erg/ili-1 homolog - sea urchin	

ORGANISM	(Lytechinus variegatus) (fragment)
DATE	#formal_name Lytechinus variegatus #common_name variegated urchin
ACCESSIONS	03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 08-Sep-1997
REFERENCE	A56646
#authors	Ol, S.; Chen, Z.Q.; Papas, T.S.; Lautenberger, J.A.
#journal	Sea Seq. (1992) 3:127-130
#title	The sea urchin erg homolog defines a highly conserved erg-specific domain.
#accession	A56646
#status	preliminary
##molecule_type	DNA
##residues	1-173 ##label Q1A
##cross-references	GB:M81067; NID:g161310; PID:g161311
CLASSIFICATION	#superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology
KEYWORDS	DNA binding; nucleus; transcription factor
FEATURE	
6-84	#domain ets DNA-binding domain homology #label ETS
SUMMARY	#length 173 #checksum 5090
Query Match	29.2%; Score 959; DB 2; Length 173;
Best Local Similarity	73.3%; Pred. No. 6.95e-171;
Matches 129; Conservative	27; Mismatches 12; Indels 8; Gaps 6;
Db	1 GSGOIQMOLFLELLSDSSNSNCITWGTNGEFGKATDPDEVARRWGRKSKPNNYDKLSR 60
Oy	280 SGOIQMOLFLELLSDSSNSNCITWGTNGEFGKATDPDEVARRWGRKSKPNNYDKLSR 339
Db	61 ARIYYDKNIMTKVHGKRAYKFDFAGLAQMOPVADPSKTRYSGLTYLPCH--PTK 118
Oy	340 ARIYYDKNIMTKVHGKRAYKFDFAGLAQMOPVADPSKTRYSGLTYLPCH--PTK 339
Db	119 LNFVGT-PIN-ESTNASLFSHSSWSPTGNTIVSGVHTPHHASMSSHTCTY 172
Oy	400 MNFVAPHPALPVTSSSEFFAAPNPWNSPTGC-TYPNTRL--P-AAHMPSHLCTTY 451
RESULT	9
ENTRY	S51226 #type fragment
TITLE	transcription factor erg/Flt-1 homolog - polychaete (Nereis diversicolor) (fragment)
ORGANISM	#formal_name Nereis diversicolor #common_name sandworm
DATE	15-Jul-1995 #sequence_revision 23-May-1997 #text_change 30-May-1997
ACCESSIONS	S51226
REFERENCE	S51226
#authors	Lelievre-Chotteau, A.; Laudet, V.; Flourens, A.; Begue, A.; Leprince, D.; Fontaine, F.
#journal	FEBS Lett. (1994) 354:62-66
#title	Identification of two ets related genes in a marine worm, the polychaete annelid Nereis diversicolor.
#accession	S51226
#status	preliminary
##molecule_type	DNA
##residues	1-179 ##label LEL
CLASSIFICATION	#superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology
KEYWORDS	DNA binding; nucleus; transcription factor
FEATURE	
7-85	#domain ets DNA-binding domain homology #label ETS
SUMMARY	#length 179 #checksum 4104
Query Match	27.2%; Score 895; DB 2; Length 179;
Best Local Similarity	67.6%; Pred. No. 3.51e-157;
Matches 121; Conservative	32; Mismatches 17; Indels 9; Gaps 8;
Db	1 GSGOIQMOLFLELLSDSSNSHITWEGNGEFGKATDPDEVARRWGRKSKPNNYDKLS 60
Oy	279 GSGOIQMOLFLELLSDSSNSHITWEGNGEFGKATDPDEVARRWGRKSKPNNYDKLS 338

Db	61	RALRYYYDKNIMTKYHGKRYAYKCFDPAGLAQAAOPSTDPAAKYQODL	LMGGYH	HTS	118
Oy	339	RALRYYYDKNIMTKYHGKRYAYKCFDPAGLAQAAOPSTDPAAKYQODL	LMGGYH	HTS	118
Db	119	KNLMAAHP-MASSASGCFPPAPYMSLVSGNLXPINSHHMSHHGCHSSHTAGSY	176		
Oy	399	KNMFVAHPHPALPYVSSSFFAAPNPWNSTPGC	176		
RESULT	10				
ENTRY		S37616	#type complete		
TITLE		transcription factor clyg - fruit fly (Drosophila melanogaster)			
ALTERNATE_NAMES		transcription factor ets-97D			
ORGANISM		#formal_name Drosophila melanogaster			
DATE		13-Jan-1995 #sequence_revision 23-May-1997 #text_change 18-Sep-1998			
ACCESSIONS		S37616, S28822			
REFERENCE		S37616			
#authors		The, S.M.; Xie, X.; Smyth, F.; Papas, T.S.; Watson, D.K.; Schulz, R.A.			
#journal		Oncogene (1992) 7:2471-2478			
#file		Molecular characterization and structural organization of D-clyg, an ets proto-oncogene-related gene of Drosophila.			
#accession		S37616			
#status		preliminary			
#molecule_type		DNA			
#residues		1-464 #label THE			
#cross-references		EMBL:X68259; NID:g7942; PID:g7943			
REFERENCE		S28819			
#authors		Chen, T.; Bunting, M.; Karim, F.D.; Thummel, C.S.			
#journal		Dev. Biol. (1992) 151:176-191			
#file		Isolation and characterization of five Drosophila genes that encode an ets-related DNA binding domain.			
#accession		S28822			
#molecule_type		RNA			
#residues		298-449 #label CHE			
#cross-references		EMBL:M88477; NID:g157189; PID:g552088			
#note		the authors translated the codon AGC for residue 302 as Thr			
GENETICS					
#gene		FlyBase:ETs97D			
#cross-references		FlyBase:FBgn0004510			
#introns		22/1; 60/2; 83/3; 322/1			
CLASSIFICATION		#superfamily transcription factor elg; ets DNA-binding domain			
KEYWORDS		homology; ets RII regulatory region homology			
FEATURE		DNA binding; nucleus; transcription factor			
190-263		#domain ets RII regulatory region homology #label ERR\			
348-426		#domain ets DNA-binding domain homology #label ETS			
SUMMARY		#length 464 #molecule-weight 52658 #checksum 2714			
Query Match		20.18; Score 660; DB 1; Length 464;			
Best Local Similarity		39.68; Pred. No. 2,17e-107;			
Matches		97; Conservative 62; Mismatches 74; Indels 12; Gaps 11;			
Db	192	EQLRLKIPAAANMTAAHYVYVLEMAVKGQELVGINSDWC-MNGQELCAWTHEEFNOKL	250		
Oy	121	NERRVAVPADPTLMSDHWQMLEMAVKEGLDPLDILLONIDKELCKTRDF-QRL	179		
Db	251	-PRDGNIFWTHLQLEK--CNFV-SVYHRAKEE-QRRKQPRIMANSISTN-SGGSL	304		
Oy	180	TPSYNADILLSHYLRERGCATFIFPNTSYVPEATQITTRPDLPEQARSATWSHP	239		
Db	305	LEGRIRKSKYQSV-KSSD-SVE-STTSSNAPSNTTGGSGNGOVOLMQLLELTDCER	361		
Oy	240	TOSKAPQSSSTYPKTEDORPOLDPYOILPTPS-SRLANGSGOIQLOMFLLELSDSN	298		
Db	362	TVYIEWVNGEGEKLDDPDAVARLMEGKKRPMANVEKLSRALRYYYDGMISVSGKRF	421		
Oy	299	SNCTIATWEGNTEGEMKMDDEVARRMGERKSKPNMNDKLSRALRYYYDKNIMTKYHGKRY	451		
Db	422	AYKFD 426			

[illegible]



```

ALTERNATE_NAMES      GABPA: nuclear respiratory factor-2 alpha chain;
                      transcription factor EATF1 60K chain
ORGANISM              #formal_name Homo sapiens #common_name man
DATE                  21-Jan-1994 #sequence_revision 23-May-1997 #text_change
                      05-Jun-1998
ACCESSIONS            A48146; A46303; A55903
REFERENCE             A48146
#authors              Watanabe, H.; Sawade, J.; Yano, K.; Yamaguchi, K.; Goto, M.;
#journal              Handa, H. Biol. (1993) 13:1385-1391
#title                cDNA cloning of transcription factor EATF1 subunits with Ets
                      and notch motifs.
#cross-references     MIMD:93180783
#accession            A48146
##status              not compared with conceptual translation
##molecule_type      mRNA; protein
##residues            1-454 ##label WANT
#note                 parts of this sequence were determined by protein
                      sequencing
#date                  1993
REFERENCE             A46303
#authors              Vitbasius, J.V.; Vitbasius, C.A.; Scarpulla, R.C.
#journal              Genes Dev. (1993) 7:380-392
#title                Identity of GABP with NRF-2, a multisubunit activator of
                      cytochrome oxidase expression, reveals a cellular role for
                      an ETS domain activator of viral promoters.
#accession            A46303
##molecule_type      protein
##residues            237-248; 429-448 ##label VIR
REFERENCE             A55903
#authors              Carcot, P.; Pastier, D.; Lacorte, J.M.; Mangeney, M.; Zannis
                      V.I.; Chambaz, J.
#journal              Biochemistry (1994) 33:12139-12148
#title                Purification and characterization of nuclear factors binding
                      to the negative regulatory element D of human
                      apolipoprotein A-II promoter: a negative regulatory effect
                      is reversed by GABP, an ets-related protein.
#accession            A55903
##molecule_type      protein
##residues            26-34, 'A', 36, 'X', 38, 'X', 40-46; 151-165; 167-174; 350-359
                      ##label CAR
GENETICS              GDB:GABPA; EATF1-60; EATF1A
#gene                 ##cross-references GDB:138476; OMIM:600609
#map_position          21q21-21q22.1
COMPLEX               GA-binding protein is a heterotetramer of two alpha and two
                      beta-type chains.
FUNCTION              #description    a transcription factor that binds (via the alpha chain) to
                      GA-rich promoters
                      #pathway        known to promote transcription of apolipoprotein A-II,
                      cytochrom c oxidase chain IV, ATPase beta chain, and
                      adenovirus E4 genes
CLASSIFICATION        #superfamily transcription factor elg; ets DNA-binding domain
                      homology; ets RII regulatory region homology
                      DNA binding; nucleus; transcription factor
KEYWORDS              #domain ets RII regulatory region homology #label ERR\
                      #domain ets DNA-binding domain homology #label ETS
FEATURE               174-245
                      322-400
SUMMARY               #length 454 #molecular_weight 51311 #checksum 3079
Query Match          19.8%; Score 649; DB 1; Length 454;
Best Local Similarity 41.6%; Pred. No. 4,38e+105;
Matches 101; Conservative 60; Mismatches 62; Indels 20; Gaps 15;
Db 176 EOEKGLIPDIOWNSIDVOVLHVVVWVKESMTDIDLTL-NISGEILCSLNOEDFFQY 234
:: : : : | | | | | : : : : | : : : : | : : : : | : : : : | : : : :
QY 121 NERVIVPADPTLSTSHVROMEMAWEXEGLPDVDILLFONIDGKEICKMKTKDF-QRL 179
:: : : : | | | | | : : : : | : : : : | : : : : | : : : : | : : : :
Db 235 -PR--GETLMSHELRL-K---VYLAAQE--QOMNE-IYTI-DQP-VOLIPAS-VQSARF 281
| : : : | | | | | : : : : | : : : : | : : : : | : : : : | : : : :
QY 180 TPSTNADLLSLHMYLBERGATFFFTPTSYVEPATQITTRDPDLPEQARRRSAMTSHSP 239

```

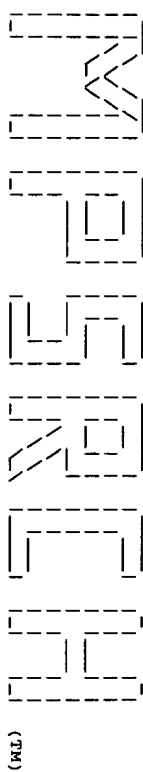
Db	282	TTIKIVN--SSAKAAKQ--BAPRISGDRSSP--GNPFGN--NGOIOLMOQFLLELTDDCAR	336
Qy	240	TSKATQPSSTSYVPKTEDRPQIDLPYQILGPTSSRLANPESGGIOLMOQFLLELTSDSNS	299
Db	337	DCISWVGDEGEFKLNPPELVACKWGRKKNPTNNYKELSRALRYYYDGDGICKVQGRFV	396
Qy	300	NCITWGTGTGERKMDPDDEVARRWGRKSKPNMNYDKLSALALYYDYDKNMTKTVHGKRYA	359
Db	397	YKF 399	
Qy	360	YKF 362	
RESULT	13		
ENTRY	138739	#type complete	
TITLE	nuclear respiratory factor-2 subunit alpha - human		
ORGANISM	#formal_name Homo sapiens #common_name man		
DATE	29-May-1998 #sequence_revision 29-May-1998 #text_change 10-Jul-1998		
ACCESSIONS	138739		
REFERENCE	138739		
#authors	Gugneja, S.; Virbasius, J.V.; Scarpulla, R.C.		
#journal	Mol. Cell. Biol. (1995) 15:102-111		
#title	Four structurally distinct, non-DNA-binding subunits of human nuclear respiratory factor 2 share a conserved transcriptional activation domain.		
#cross-references	MUID:95097980		
#accession	138739		
#status	preliminary; translated from GR/EMBL/DBJ		
#molecule_type	mRNA		
#residues	1-454 #label RES		
#cross-references	EMBL:U03044; NID:9531892; PID:9531893		
CLASSIFICATION	#superfamily transcription factor elg; ets DNA-binding domain; homology; ets RII regulatory region homology		
FEATURE			
174-245	#domain ets RII regulatory region homology #label EKR		
SUMMARY	#length 454 #molecular_weight 51408 #checksum 3243		
Query Match	19.5%; Score 640; DB 2; Length 454;		
Best Local Similarity	40.7%; Pred. No. 3,35e-103;		
Matches	99; Conservative 61; Mismatches 63; Indels 20; Gaps 15;		
Db	176	EOERLGIPIPIOWSTDOYLHWVYWMKEFSMTDIDLTLL-NISGRELCSLQDEDFQRY	234
Qy	121	NERRVIVPADPRLTMSDHWROWLEMAVKEKGLPDVDIILFQONIDKELCKMTKDF-QRL	179
Db	235	-RR--CELLMSHELAR-K---VYLAQQ--QQMNE-IYVI-DQP-VQIIPAS-VQSATP	281
Qy	180	FTSYNADILLSHLYREGAGTFIFENTSYVPATORTITRPLPEQARRSAATSHSP	239
Db	282	TTIKIVINRCAKA-AKQV-BAPRISGDRSSP--GNPFGN--NGOIOLMOQFLLELTDDCAR	336
Qy	240	TSKATQPSSTSYVPKTEDRPQIDLPYQILGPTSSRLANPESGGIOLMOQFLLELTSDSNS	299
Db	337	DCISWVGDEGEFKLNPPELVACKWGRKKNPTNNYKELSRALRYYYDGDGICKVQGRFV	396
Qy	300	NCITWGTGTGERKMDPDDEVARRWGRKSKPNMNYDKLSALALYYDYDKNMTKTVHGKRYA	359
Db	397	YKF 399	
Qy	360	YKF 362	
RESULT	14		
ENTRY	528823	#type fragment	
TITLE	transcription factor ets-21C - fruit fly (Drosophila melanogaster) (fragment)		
ALTERNATE_NAMES	transforming protein ets-6		
ORGANISM	#formal_name Drosophila melanogaster		
DATE	31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997		
ACCESSIONS	528823		
REFERENCE	528819		

Job time : 68 secs.

```
#authors      Chen, T.; Bunting, M.; Karim, F. D.; Thummel, C.S.
#journal      Dev. Biol. (1992) 151:176-191
#title        Isolation and characterization of five Drosophila genes that
               encode an ets-related DNA binding domain.
#accession    S28823
               #molecule_type DNA
               #residues      1-103 #label CHE
               ##cross-references EMBL:M88475
GENETICS
#gene         Ets21C
               #cross-references FlyBase:FBgn0005660
               #introns      64/2
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain
               #homology; ets RII regulatory region homology
KEYWORDS       DNA binding; nucleus; transcription factor
FEATURE        8-86
SUMMARY        #domain ets DNA-binding domain homology #label ETS
               #length 103 #checksum 5825
Query Match    19.0%; Score 623; DB 2; Length 103;
Best Local Similarity 80.4%; Pred. No. 1.19e-99;
Matches 78; Conservative 13; Mismatches 6; Indels 0; Gaps 0;
Db 2 SGGQIQLMQFLLELADSSNANAIWEGSGEPRLDPDEVARRGGERKAKPMNYDKLS 61
QY 279 GSGQIQLMQFLLELSDSSNSNCITWGTNGEFTKMTDPDEVARRGGERKSKPMNYDKLS 338
Db 62 RALRYYYDKNIMTKVHGKRYAYKFDFFGLMAACQACA 98
QY 339 RALRYYYDKNIMTKVHGKRYAYKFDFFGLMAALQPHR 375

RESULT 15
ENTRY       S24300 #type complete
TITLE       transforming protein (elg) - fruit fly (Drosophila
               melanogaster)
ORGANISM    #formal_name Drosophila melanogaster
DATE        13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
               16-Feb-1997
ACCESSIONS  S24300
REFERENCE   S24300
#authors    Pribyl, L.J.; Watson, D.K.; Schulz, R.A.; Papas, T.S.
#journal     Oncogene (1991) 6:1175-1183
#title       D-clg, a member of the Drosophila ets gene family: sequence,
               expression and evolutionary comparison.
#accession   S24300
               #status      preliminary
               ##molecule_type mRNA
               ##residues    1-211 #label PRI
               ##cross-references EMBL:X58481
GENETICS
#gene        FlyBase:Ets97D
               #cross-references FlyBase:FBgn0004510
CLASSIFICATION #superfamily ets DNA-binding domain homology
KEYWORDS       DNA binding
FEATURE        95-173
SUMMARY        #domain ets DNA-binding domain homology #label ETS
               #length 211 #molecular_weight 23603 #checksum 6044
Query Match    15.0%; Score 493; DB 2; Length 211;
Best Local Similarity 59.8%; Pred. No. 8.80e-73;
Matches 58; Conservative 24; Mismatches 14; Indels 1; Gaps 1;
Db 77 MNPSYTTIGSGNGQVOIQMFLLELIDCEHTDVIENWGTGGEFTKMTDPDEVARRLWGER 136
QY 268 LGPTS-SRLANPGSQIQMFLLELSDSSNSNCITWGTNGEFTKMTDPDEVARRWGER 326
Db 137 KKKPMANTEKLSRALRYYYDGMISKVSGKRPAKKFD 173
QY 327 KSKPMNMYDKLSRALRYYYDKNIMTKVHGKRYAYKFD 363

Search completed: Wed Apr 7 09:33:14 1999
```



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MSrch\_p protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 7 09:33:30 1999; Maspar time 13.48 Seconds  
898.017 Million cell updates/sec

Output not generated.

Title: >US-08-878-177-2  
Description: (1-451) from US08878177.pep  
Perfect Score: 3286  
Sequence: 1 MASTIKKALSVSEDSLFE.....IYPTNRLPAHMSHGLTY 451

Scoring table: PAM 150  
Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot35  
1:swissprot

Statistics: Mean 49.266; Variance 82.055; Scale 0.600

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2991	91.0	462	1	ERG_HUMAN TRANSFORMING PROTEIN E	0.00e+00
2	2348	71.5	452	1	FLI1_HUMAN FLI1 ONCOGENE (ERGB T	0.00e+00
3	2339	71.2	452	1	FLI1_MOUSE RETROVIRAL INTEGRATION	0.00e+00
4	2305	70.1	453	1	FLI1_XENLA RETROVIRAL INTEGRATION	0.00e+00
5	1535	46.7	272	1	ERG_MOUSE TRANSFORMING PROTEIN E	0.00e+00
6	959	29.2	173	1	ERG_LYTA ERG PROTEIN HOMOLOG (F	2.43e-197
7	650	20.1	464	1	ERG_DROME ERG-BINDING PROTEIN D-	5.29e-124
8	653	19.9	454	1	GABA_MOUSE GA BINDING PROTEIN ALP	2.61e-122
9	649	19.8	454	1	GABA_HUMAN GA BINDING PROTEIN ALP	2.61e-122
10	623	19.0	102	1	ETS6_DROME DNA-BINDING PROTEIN D-	4.54e-115
11	492	15.0	268	1	ET1B_XENLA C-ETS-1B PROTEIN (XEL-	7.87e-84
12	492	15.0	438	1	ET1A_XENLA C-ETS-1A PROTEIN (XEL-	7.87e-84
13	490	14.9	440	1	ETSL_MOUSE C-ETS-1 PROTEIN (P54).	2.33e-83
14	490	14.9	441	1	ETSL_HUMAN C-ETS-1 PROTEIN (P54).	2.33e-83
15	490	14.9	441	1	ETSL_RAT C-ETS-1 PROTEIN (P54).	2.33e-83
16	486	14.8	441	1	ETSA_CHICK TRANSFORMING PROTEIN P	2.04e-82
17	487	14.8	472	1	ET2A_XENLA C-ETS-2A PROTEIN.	2.04e-82
18	486	14.8	485	1	ETSB_CHICK TRANSFORMING PROTEIN P	2.04e-82
19	485	14.8	623	1	PMT1_DROME ETS-LIKE PROTEIN POINT	3.50e-82
20	485	14.8	718	1	PMT2_DROME ETS-LIKE PROTEIN POINT	3.50e-82
21	483	14.7	468	1	ETS2_MOUSE C-ETS-2 PROTEIN.	1.03e-81
22	483	14.7	469	1	ETS2_HUMAN C-ETS-2 PROTEIN.	1.03e-81
23	484	14.7	472	1	ET2B_XENLA C-ETS-2B PROTEIN.	6.01e-82

24	482	14.7	669	1	MYB_AVILE P135-GAG-MYB-ETS TRANS	1.78e-81
25	480	14.6	479	1	ETS2_CHICK C-ETS-2 PROTEIN.	5.24e-81
26	474	14.4	110	1	ETS2_LYTA C-ETS-2 PROTEIN (FRAGM	1.34e-79
27	469	14.3	548	1	ERG_HUMAN ETS-DOMAIN TRANSCRIPT	2.00e-78
28	469	14.3	551	1	ERG_MOUSE ETS-DOMAIN TRANSCRIPT	2.00e-78
29	459	14.0	64	1	ETS3_DROME DNA-BINDING PROTEIN D-	4.39e-76
30	460	14.0	477	1	ETV1_MOUSE ETS1 PROTEIN (ETS TRAN	2.56e-76
31	460	14.0	477	1	ETV1_HUMAN ETS1 PROTEIN (ETS TRAN	2.56e-76
32	456	13.9	250	1	ETV3_HUMAN ETS-RELATED PROTEIN PE	2.21e-75
33	454	13.8	510	1	ERM_HUMAN ETS-RELATED PROTEIN ER	6.47e-75
34	443	13.5	335	1	ETV2_MOUSE ETS-RELATED PROTEIN 71	2.38e-72
35	444	13.5	341	1	ETV2_HUMAN ETS-RELATED PROTEIN 71	2.39e-72
36	443	13.5	551	1	ETV4_HUMAN ADENOVIRUS E1A ENHANCE	2.38e-72
37	443	13.5	555	1	ETV4_MOUSE POLYOMAVIRUS ENHANCER	2.38e-72
38	437	13.3	428	1	ELK1_HUMAN ETS-DOMAIN PROTEIN ELK	5.91e-71
39	431	13.1	429	1	ELK1_MOUSE ETS-DOMAIN PROTEIN ELK	1.46e-69
40	429	13.1	430	1	SAPA_MOUSE SERUM RESPONSE FACTOR	4.26e-69
41	422	12.8	407	1	ELK3_HUMAN ETS-DOMAIN PROTEIN ELK	1.78e-67
42	420	12.8	409	1	ELK3_MOUSE ETS-DOMAIN PROTEIN ELK	1.58e-67
43	418	12.7	405	1	SAPB_HUMAN SERUM RESPONSE FACTOR	1.50e-66
44	418	12.7	431	1	SAPA_HUMAN SERUM RESPONSE FACTOR	1.50e-66
45	372	11.3	619	1	ELF1_HUMAN ETS-RELATED TRANSCRIPT	5.39e-56

## ALIGNMENTS

RESULT	1	STANDARD;	PRT;	462 AA.
ID	ERG_HUMAN			
AC	P13508;			
DT	01-JUL-1989 (REL. 11, CREATED)			
DT	01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	TRANSFORMING PROTEIN ERG (ERG-2 / ERG-1).			
GN	ERG.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 87263429.			
RA	RAO V.N., PAPAS T.S., SHRAM E., REDDY P.;			
RL	SCIENCE 237:635-639(1987).			
RN	[2]			
RP	SEQUENCE OF 100-462 FROM N.A.			
RX	MEDLINE; 87317608.			
RA	REDDY E.S.P., RAO V.N., PAPAS T.S.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 84:6131-6135(1987).			
RN	[3]			
RP	CHROMOSOMAL TRANSLOCATION.			
RX	MEDLINE; 94356859.			
RA	DUNN T., PRAISSMAN L., HAGAG N., VIOLA M.V.;			
RL	CANCER GENET. CYTOGENET. 76:19-22(1994).			
RN	[4]			
RP	CHROMOSOMAL TRANSLOCATION.			
RX	MEDLINE; 94243799.			
RA	ICHIKAWA H., SHIMIZU K., HAYASHI Y., OHKI M.;			
RL	CANCER RES. 54:2865-2868(1994).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED			
CC	BY A CHROMOSOMAL TRANSLOCATION T(16;21)(p11.022) WHICH INVOLVES			
CC	ERG AND FUS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
DR	EMBL; M17254; G182187; -.			
DR	EMBL; M21535; G182185; -.			
DR	PIR; A29515; TVHUG.			
DR	TRANSFAC; T00265; -.			
DR	TRANSFAC; T00266; -.			
DR	MTM; 165080; -.			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE; PS50061; ETS_DOMAIN_3; 1.			
KW	NUCLEAR PROTEIN; TRANSFORMING PROTEIN; DNA-BINDING;			
KW	ALTERNATIVE INITIATION; PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION.			



QY 418 FAAPNPVNSPTGGIYPTNT---RLPAAHMPSHLGTY 451

RESULT 3 STANDARD: PRT: 452 AA.

ID FULL\_MOUSE

AC P26323;

DT 01-MAY-1992 (REL. 22, CREATED)

DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE RETROVIRAL INTEGRATION SITE PROTEIN FLI-1.

GN FLI1 OR FLI-1.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; RODENTIA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-SPLEEN;

RX MEDLINE; 91257578.

RA BEN-DAVID Y., GIDDENS E.B., LETWIN K., BERNSTEIN A.;

RA GENES DEV. 5:908-918(1991).

[2]

CHARACTERIZATION

RE MEDLINE; 93275657.

RX ZHANG L., LEMARCHANDEL V., ROMEO P.-H., BEN-DAVID Y., GREER P.,

RA BERNSTEIN A.;

RL ONCOGENE 8:1621-1630(1993).

CC -1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES

CC -1- THE DNA SEQUENCE 5'C[CA]GGAAGT-3'.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH HEMATOPOIETIC AND

CC -1- NONHEMATOPOIETIC TISSUES.

CC -1- INDUCTION: SPECIFICALLY UP-REGULATED IN LEUKEMIC CLONES WITH

CC -1- F-MULV INSERTIONS UP-STREAM OF THE FLI-1 LOCUS

CC -1- DISEASE: INVOLVED IN ERYTHROLEUKEMIA INDUCTION BY FRIEND MURINE

CC -1- LEUKEMIA VIRUS (F-MULV).

CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

DR EMBL; X59421; G50975; -.

DR PIR; S17403; S17403.

DR TRANSFAC; T01408; -.

DR MGD; MGI:95554; FLI1.

DR PROSITE; PS00345; ETS\_DOMAIN\_1; 1.

DR PROSITE; PS00346; ETS\_DOMAIN\_2; 1.

DR PROSITE; PS50061; ETS\_DOMAIN\_3; 1.

KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;

KW PROTO-ONCOGENE.

FT DNA\_BIND 281 361 ETS-DOMAIN.

SO SEQUENCE 452 AA; 51002 MM; CECACDD4 CRC32;

Query Match 71.2%; Score 2339; DB 1; Length 452;

Best Local Similarity 66.7%; Pred. No. 0.00e+00;

Matches 302; Conservative 85; Mismatches 56; Indels 11; Gaps 9;

Db 1 MDGITEKALSVSDQSLFSDATYGAHAHLPKADMTASGSPDYQPHKINPLPPQDWMNQ 60

QY 1 MASTIKEALSVSEDSQSLFECAYGSP-HLAKTEMTASSSEYQGTSSKSPRVPODDWLSQ 59

Db 61 P-VARNVREYD--HMNSRESYPVDCSYSKCNKLVGGGEANPMNTNITYDEKNGPPPM 117

QY 60 PPARVTIMKCNPNQVNSRNSPDDCSYAKGKMYSSSDNMGMYGSMYEEKH-1PPPM 118

Db 118 TTNERRVIVPADPTLMTGSDHVRQMLEMAIKKEYGLMEIDTSFEQNDGKELCKMKNEDFLR 177

QY 119 TTNERRVIVPADPTLMTGSDHVRQMLEMAIKKEYGLMEIDTSFEQNDGKELCKMKNEDFLR 178

Db 178 ATSAVTEVLLSHLSLYRE-SSLAY-NTSHTDQSSRLNVEDPSYDSVRGANNNNMN 235

QY 179 LTPSYNADILLSHLYLREKATGATFEFPNTSVYPEATQRTTRPDLPYQARSAMTSYSH 238

Db 236 SGLNKSPLLGSGQTMKTEQRPQDPYQIIGTSSRLANPSSGQIQMLFLELLSDSA 295

QY 239 PLOSATOPSSS-TVPKTEDORPOLDPYQIIGTSSRLANPSSGQIQMLFLELLSDS 297

Db 296 NASCTTGEKTGERKMDPDDEVARMRGKSKPMNNNDKLSRALRYTDKNIMTKVGR 355

QY 298 NSNCITWETGNGEKKMDPDDEVARMRGKSKPMNNNDKLSRALRYTDKNIMTKVGR 357

Db 356 YAKFDFHGIQAOLQPHPTETSMKYPSDIDSYMSYHAHQKVPNPVSPSSMPTSSSF 415

QY 358 YAKFDFHGIQAOLQPHPTETSMKYPSDIDSYMSYHAHQKVPNPVSPSSMPTSSSF 417

Db 416 FGASQVTSPTAGIYPNPSVRPHNTVPSHLSGY 452

QY 418 FAAPNPVNSPTGGIYPTNT---RLPAAHMPSHLGTY 451

RESULT 4 STANDARD: PRT: 453 AA.

ID FLI1\_XENLA

AC P4157;

DT 01-FEB-1995 (REL. 31, CREATED)

DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)

DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)

DE RETROVIRAL INTEGRATION SITE PROTEIN FLI-1 HOMOLOG.

GN FLI1.

OS XENOPUS LAEVIS (AFRICAN CLAMPED FROG).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 94206844.

RA MEYER D., WOLFF C.M., STIEGLER P., SENAN F., BEFORT N.,

RA BEFORT J.J., REMY P.;

RL MECH. DEV. 44:109-121(1993).

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.

CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

DR EMBL; X66979; G505487; -.

DR TRANSFAC; T02067; -.

DR PROSITE; PS00345; ETS\_DOMAIN\_1; 1.

DR PROSITE; PS00346; ETS\_DOMAIN\_2; 1.

DR PROSITE; PS50061; ETS\_DOMAIN\_3; 1.

KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN.

FT DNA\_BIND 282 362 ETS-DOMAIN.

SO SEQUENCE 453 AA; 51015 MM; DEB5A908 CRC32;

Query Match 70.1%; Score 2305; DB 1; Length 453;

Best Local Similarity 65.9%; Pred. No. 0.00e+00;

Matches 302; Conservative 85; Mismatches 59; Indels 12; Gaps 9;

Db 1 MDGITEKALSVSDQSLFSDATYGAASHLSKADMTASNPDIYQPHKINPLPPQDWMNQ 60

QY 1 MASTIKEALSVSEDSQSLFECAYG-SPHLAKTEMTASSSEYQGTSSKSPRVPODDWLSQ 59

Db 61 P-MRVNIREYE--HMNSRESYPVDCSYKCNKLVGGGEANPMNTNITYDEKNGPPPM 116

QY 60 PPARVTIMKCNPNQVNSRNSPDDCSYAKGKMYSSSDNMGMYGSMYEEKH-1PPPM 118

Db 117 TTNERRVIVPADPALMSODHVRQMLEMAIKKEYGLVEIDCSLFQNDGKELCKMKNEDFLR 176

QY 119 TTNERRVIVPADPTLMTGSDHVRQMLEMAIKKEYGLVDVILLFQNDGKELCKMKNEDFLR 178

Db 177 STSIYNTVLLSHLYLFDSSSLGY-NTQAHTDQSSRLNVEDPSYAVRRSGNGSMS 235

QY 179 LTPSYNADILLSHLYLREKATGATFEFPNTSVYPEATQRTTRPDLPYQARSAM-TSHS 237

Db 236 SPVTKSPMGQNVNKSQDQORSQPDYQIIGTSSRLANPSSGQIQMLFLELLSDS 295

QY 238 HPTOSKATQSSSTVPK-EDQRPOLDYQIIGTSSRLANPSSGQIQMLFLELLSDS 296

Db 296 SNASCTTGEKTGERKMDPDDEVARMRGKSKPMNNNDKLSRALRYTDKNIMTKVGR 355

QY 297 SNSNCITWETGNGEKKMDPDDEVARMRGKSKPMNNNDKLSRALRYTDKNIMTKVGR 356

Db 356 YAKFDFHGIQAOLQPHPTETSMKYPSDIDSYMSYHAHQKVPNPVSPSSMPTSSSF 415

QY 357 YAKFDFHGIQAOLQPHPTETSMKYPSDIDSYMSYHAHQKVPNPVSPSSMPTSSSF 416



QY 121 NERVIAPADPTLWSTHVRQWLEMAVKEYGLPVDILLFQNDIGKELCKTKDQDF-QRL 179  
Db 251 -PRRPGIIFWTHLOLKE--CNFV-SVYHKAEE-ORRKPORRMSANSISTN-SGGSLS 304  
QY 180 TPSYNADILLSHLYLBERGATFIFPNTSVYPEATQRTTRTPDLPYEQARSAMTSHSP 239  
Db 305 LEORIMRKSYSOV-KSSD-SVE-STTSSMNPSTTTGSGNNGOVOLQWLFLELTIDCEH 361  
QY 240 TQSRATOPSSSTVKTEDQRPQDLPYQILGFTS-SRLANPSSGOIQLOWLFLELSDSN 298  
Db 362 TDVIEWGTGEGFKLTDPPDRVARKLGEKKRKPANNEYKLSALRYDDGDMISVSGRRF 421  
QY 299 SNTCTWEGTNGEFKMTDPEVARRMGERKSKPNMNYDKLSRALRYDDKINTKYGKRY 358  
Db 422 AYKFD 426  
QY 359 AYKFD 363

ULT 8 GABA\_MOUSE STANDARD: PRT: 454 AA.

000422;  
01-DEC-1992 (REL. 24, CREATED)  
01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE GA BINDING PROTEIN ALPHA CHAIN (GABP-ALPHA SUBUNIT).  
GABPA.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 91343912.  
RA LAMARCO K., THOMPSON C.C., BYERS B.P., WALTON E.M., MCKNIGHT S.L.;  
SCIENCE 253:789-792(1991).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 320-320.  
RX MEDLINE: 98128030.  
RA BATCHELOR A.H., PIPER D.E., DE LA BROUSSE F.C., MCKNIGHT S.L.,  
WOLBERGER C.;  
SCIENCE 279:1037-1041(1998).  
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE  
RICH REPEATS (GA REPEATS).  
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- TISSUE SPECIFICITY: OBITUOUS.  
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
EMBL: M74515; G193383; -.  
PIR: A40858; A40858.  
PDB: 1AWC; 18-MAR-98.  
DR TRANSFAC: T00298; -.  
DR TRANSFAC: T01402; -.  
DR MGD; MGI:95610; GABPA.  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN; 3D-STRUCTURE.  
FT DNA\_BIND 320 400 ETS-DOMAIN.  
SQ SEQUENCE 454 AA; 51363 MW; C3BA5F76 CRC32;

Query Match 19.9% Score 653; DB 1; Length 454;  
Best Local Similarity 41.6% Pred. No. 2,61e-122;  
Matches 101; Conservative 61; Mismatches 61; Indels 20; Gaps 15;

Db 176 EOEELGIPYDPIRSTQOVHLHVVVWVWKEFMTDIDLTTL-NISGRELSLNQEDFFQRY 234  
QY 121 NERVIAPADPTLWSTHVRQWLEMAVKEYGLPVDILLFQNDIGKELCKTKDQDF-QRL 179  
Db 235 -PR-GEILWSHLELR-K--YVLASOE--QONNE-IVTI-DOP-VOIIPAS-VPPATP 281  
QY 180 TPSYNADILLSHLYLBERGATFIFPNTSVYPEATQRTTRTPDLPYEQARSAMTSHSP 239

Db 282 TTIKVIN-SSAKAKVO-RSPRISGEDRSSP-GNRGTN--NGOIQLOWLFLELTIDRDAR 336  
QY 240 TQSRATOPSSSTVKTEDQRPQDLPYQILGFTSRLANPSSGOIQLOWLFLELSDSN 299  
Db 337 DCISWGDGEGFKLNOBELVAQKWKGRKKNPTMYEKLRLRYDDGDMICKYQGRFV 396  
QY 300 NCITWEGTNGEFKMTDPEVARRMGERKSKPNMNYDKLSRALRYDDKINTKYGKRYA 359  
Db 397 YKF 399  
QY 360 YKF 362

RESULT 9 GABA\_HUMAN STANDARD: PRT: 454 AA.

AC 006546; 012939;  
01-FEB-1995 (REL. 31, CREATED)  
01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE GA BINDING PROTEIN ALPHA CHAIN (GABP-ALPHA SUBUNIT) (TRANSCRIPTION  
FACTOR EATF1-60) (NUCLEAR RESPIRATORY FACTOR-2 SUBUNIT ALPHA).  
GN EATF1A OR GABPA.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93180783.  
RA WATANABE H., SAWADA J.-I., YANO K.-I., YAMAGUCHI K., GOTO M.,  
HANDA H.;  
MOL. CELL. BIOL. 13:1385-1391(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 95097980.  
RA GUGNEVA S., VIRBASJUS J.V., SCARPUOLA R.C.;  
MOL. CELL. BIOL. 15:102-111(1995).  
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE  
RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE  
ADENOVIRUS EA GENE.  
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
DR EMBL: D13318; G286027; -.  
DR EMBL: U13044; G531893; -.  
DR TRANSFAC: T01390; -.  
DR MIM; 600609; -.  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.  
FT DNA\_BIND 320 400 ETS-DOMAIN.  
FT CONFLICT 289 290 SS -> RC (IN REF. 2).  
FT CONFLICT 440 440 A -> V (IN REF. 2).  
SQ SEQUENCE 454 AA; 51295 MW; E0EA85D4 CRC32;

Query Match 19.8% Score 649; DB 1; Length 454;  
Best Local Similarity 41.6% Pred. No. 2,42e-121;  
Matches 101; Conservative 60; Mismatches 62; Indels 20; Gaps 15;

Db 176 EOEELGIPYDPIRSTQOVHLHVVVWVWKEFMTDIDLTTL-NISGRELSLNQEDFFQRY 234  
QY 121 NERVIAPADPTLWSTHVRQWLEMAVKEYGLPVDILLFQNDIGKELCKTKDQDF-QRL 179  
Db 235 -PR-GEILWSHLELR-K--YVLASOE--QONNE-IVTI-DOP-VOIIPAS-VQSATP 281  
QY 180 TPSYNADILLSHLYLBERGATFIFPNTSVYPEATQRTTRTPDLPYEQARSAMTSHSP 239  
Db 282 TTIKVIN-SSAKAKVO-RAPRISGEDRSSP-GNRGTN--NGOIQLOWLFLELTIDRDAR 336  
QY 240 TQSRATOPSSSTVKTEDQRPQDLPYQILGFTSRLANPSSGOIQLOWLFLELSDSN 299  
Db 337 DCISWGDGEGFKLNOBELVAQKWKGRKKNPTMYEKLRLRYDDGDMICKYQGRFV 396

ID	ETIA_XENLA	STANDARD	PRT	438 AA.
AC	P18755.			
DT	01-NOV-1990 (REL. 16, CREATED)			
DT	01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	C-ETS-1A PROTEIN.			
GN	ETS-1A.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EVARUOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: AMPHIBIA. ANURA.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=OVARY:			
RX	MEDLINE: 90384849.			
RA	STIEGLER P., WOLFF C.M., BALZINGER M., HIRZLIN J., SEUWEN F., MEYER D., GHISDAEL J., STEHELIN D., BEFORT N., REMY P., NUCLEIC ACIDS RES. 18:5298-5298(1990).			
RL	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- ALTERNATIVE PRODUCTS: GIVES RISE TO C-ETS-1A' LACKING THE FIRST 203 AMINO ACIDS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
CC	EMBL: X52692; G64615; -.			
DR	PIR: S11225; S11225.			
DR	TRANSFAC: T02040; -.			
DR	PROSITE: PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE: PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE: PS50061; ETS_DOMAIN_3; 1.			
DR	DNA-BINDING; NUCLEAR PROTEIN; PHOSPHORYLATION.			
FT	VASAPLIC 1 203 MISSING (IN C-ETS-1A').			
FT	DNA_BIND 332 412 ETS-DOMAIN.			
FT	MOD_RES 153 153 PHOSPHORYLATION (BY CAM-KINASE) (POTENTIAL).			
FT	MOD_RES 279 279 PHOSPHORYLATION (BY CAM-KINASE) (POTENTIAL).			
FT	FT FT			
FT	MOD_RES 109 109 PHOSPHORYLATION (BY CAM-KINASE) (POTENTIAL).			
FT	FT			
SO	SEQUENCE 268 AA; 30772 MW; 6EF0772C CnC32;			
Query Match	15.0%; Score 492; DB 1; Length 268;			
Best Local Similarity	59.0%; Pred. No. 7,87e-84;			
Matches	62; Conservative 20; Mismatches 23; Indels 0; Gaps 0;			
Db	137 DRAELNCKDPVTPAALAGTSGSPIDLOMFOLELLTDKSCQSISTGTGCMERKLSDDP 196			
Qy	258 ORPOLDPQIOLGPTSSSRILANGSGOILOMFOLELLTDSSNSNCITWETNGEFGKMTDPD 317			
Db	197 EVARRWGRKRNKPKMYNEKLSRGIRYYDDNIHKTGKRYVRE 241			
Qy	318 EVARRWGRKRNKPKMYNEKLSRGIRYYDDNIHKTGKRYVRE 362			
RESULT	12			
ID	ETIA_XENLA	STANDARD	PRT	438 AA.
AC	P18755.			
DT	01-NOV-1990 (REL. 16, CREATED)			
DT	01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	C-ETS-1A PROTEIN.			
GN	ETS-1A.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EVARUOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: AMPHIBIA. ANURA.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=OVARY:			
RX	MEDLINE: 90384849.			
RA	STIEGLER P., WOLFF C.M., BALZINGER M., HIRZLIN J., SEUWEN F., MEYER D., GHISDAEL J., STEHELIN D., BEFORT N., REMY P., NUCLEIC ACIDS RES. 18:5298-5298(1990).			
RL	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- ALTERNATIVE PRODUCTS: GIVES RISE TO C-ETS-1A' LACKING THE FIRST 203 AMINO ACIDS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
CC	EMBL: X52692; G64615; -.			
DR	PIR: S11225; S11225.			
DR	TRANSFAC: T02040; -.			
DR	PROSITE: PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE: PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE: PS50061; ETS_DOMAIN_3; 1.			
DR	DNA-BINDING; NUCLEAR PROTEIN; PHOSPHORYLATION.			
FT	VASAPLIC 1 203 MISSING (IN C-ETS-1A').			
FT	DNA_BIND 332 412 ETS-DOMAIN.			
FT	MOD_RES 153 153 PHOSPHORYLATION (BY CAM-KINASE) (POTENTIAL).			
FT	MOD_RES 279 279 PHOSPHORYLATION (BY CAM-KINASE) (POTENTIAL).			
FT	FT FT			
FT	MOD_RES 109 109 PHOSPHORYLATION (BY CAM-KINASE) (POTENTIAL).			
FT	FT			
SO	SEQUENCE 268 AA; 30772 MW; 6EF0772C CnC32;			
Query Match	15.0%; Score 492; DB 1; Length 438;			
Best Local Similarity	59.0%; Pred. No. 7,87e-84;			
Matches	62; Conservative 20; Mismatches 23; Indels 0; Gaps 0;			
Db	307 DRAELNCKDPVTPAALAGTSGSPIDLOMFOLELLTDKSCQSISTGTGCMERKLSDDP 366			
Qy	258 ORPOLDPQIOLGPTSSSRILANGSGOILOMFOLELLTDSSNSNCITWETNGEFGKMTDPD 317			
Db	367 EVARRWGRKRNKPKMYNEKLSRGIRYYDDNIHKTGKRYVRE 411			



||||| :||:|||||:|||||:| ||| :|  
OY 318 EVARWGERSKPMNNDKLSRALRYDDKNIMTKVGRKRYAKF 362  
RESULT 13  
ID ETS1\_MOUSE STANDARD; PRT; 440 AA.  
AC P27577; 061403;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE C-ETS-1 PROTEIN (P54).  
GN ETS1 OR ETS-1.  
OS MUS MUSCULUS (MURINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WATSON D.K., SETH A., SMITH F.E., SCHWEINEST C.W., PAPAS T.S.;  
RL (IN) ONCOGENESIS, PAPAS T.S., ED., PP. 221-232, GULF PUBLISHING  
COMPANY, HOUSTON, (1990).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE-THYMUS;  
RX MEDLINE: 90299137.  
RA GUNTER C.V., NYE J.A., BRYNER R.S., GRAVES B.J.;  
RL GENES DEV. 4:667-679(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE-THYMUS;  
RX MEDLINE: 90370376.  
RA CHEN J.H.;  
RL ONCOGENE RES. 5:277-285(1990).  
RN [4]  
RP STRUCTURE BY NMR OF 332-415.  
RX MEDLINE: 96176767.  
RA DONALDSON L.W., PETERSEN J.M., GRAVES B.J., MCINTOSH L.P.;  
RL EMBO J. 15:125-134(1996).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
DR EMBL: M58483; G193192; -.  
DR EMBL: X55953; G50871; -.  
DR EMBL: X55787; G296023; -.  
DR PIR: A30487; A30487.  
DR PDB: 1ETC; 29-JAN-96.  
DR TRANSFAC: T00111; -.  
DR MGD: MGI:93455; ETS1.  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
KW PROTO-ONCOGENE; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING;  
KW PHOSPHORYLATION; 3D-STRUCTURE.  
ET DNA\_BIND 335 415 ETS-DOMAIN.  
ET CONFLICT 28 28 D-> E (IN REF. 1 AND 3).  
ET CONFLICT 37 37 L-> S (IN REF. 3).  
ET CONFLICT 51 52 AT-> SY (IN REF. 3).  
ET CONFLICT 55 55 G-> P (IN REF. 1).  
ET CONFLICT 63 63 L-> R (IN REF. 3).  
ET CONFLICT 74 74 E-> D (IN REF. 3).  
ET CONFLICT 96 96 Q-> H (IN REF. 3).  
ET CONFLICT 105 105 L-> V (IN REF. 3).  
ET CONFLICT 157 157 D-> V (IN REF. 3).  
ET CONFLICT 211 211 Q-> R (IN REF. 3).  
ET CONFLICT 217 217 D-> E (IN REF. 3).  
ET CONFLICT 225 225 A-> R (IN REF. 3).  
ET CONFLICT 234 234 D-> N (IN REF. 3).  
ET CONFLICT 360 360 G-> C (IN REF. 3).  
ET CONFLICT 383 383 K-> S (IN REF. 3).  
ET CONFLICT 392 392 G-> A (IN REF. 3).  
ET CONFLICT 408 408 KR-> NA (IN REF. 3).  
ET CONFLICT 413 413 R-> A (IN REF. 3).  
SQ SEQUENCE 440 AA; 50201 MM; 18E90F01 CRC32;

Query Match 14.9%; Score 490; DB 1; Length 440;  
Best Local Similarity 59.0%; Pred. No. 2,33e-83;  
Matches 62; Conservative 20; Mismatches 23; Indels 0; Gaps 0;  
Db 310 DRADLNDRPYPAALAGTGGSPGIQMOFLELLDLSQSGSFISWTGDCWEKRLSDPD 369  
OY 258 QRPQDPYQIIGPTSSSLANGSGQIQLOMFLLELSDSSNSNCITWEGTNGEERKMDPD 317  
Db 370 EVARWGERSKPMNNDKLSRALRYDDKNIMTKVGRKRYAKF 414  
OY 318 EVARWGERSKPMNNDKLSRALRYDDKNIMTKVGRKRYAKF 362  
RESULT 14  
ID ETS1\_HUMAN STANDARD; PRT; 441 AA.  
AC P14921;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE C-ETS-1 PROTEIN (P54) (C-ETS-1A AND C-ETS-1B).  
GN ETS1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA REDDY E.S.P., RAO V.N.;  
RX MEDLINE: 89083219.  
RA ONCOGENE RES. 3:239-246(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 89042086.  
RA WATSON D.K., MCWILLIAMS M.J., LAPIS P., LAUTENBERGER J.A.,  
RA SCHWEINEST C.W., PAPAS T.S.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 85:7862-7866(1988).  
RN [3]  
RP STRUCTURE BY NMR OF 320-415.  
RX MEDLINE: 96097120.  
RA WERNER M.H., CLORE G.M., FISHER C.L., FISHER R.J., TRINH L.,  
RA SHILOACH J., GRONENBORN A.M.;  
RL CELL 83:761-771(1995).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- PTM: ISOFORMS OF ETS-1 ARE ALSO GENERATED BY DIFFERENTIAL  
PHOSPHORYLATION.  
CC -1- DISEASE: ETS IS RESPONSIBLE FOR ERTHTROBLAST AND FIBROBLAST  
TRANSFORMATION. THE JUXTAPPOSITION OF THE INTERFERON AND C-ETS-1  
PROTO-ONCOGENE MAY BE INVOLVED IN THE PATHOGENESIS OF HUMAN  
MONOCYTIC LEUKEMIA.  
CC -1- ALTERNATIVE PRODUCTS: ISOFORMS OF ETS-1 CAN BE GENERATED BY  
ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
DR EMBL: X14798; G29882; -.  
DR EMBL: X14798; G296642; -.  
DR EMBL: J04101; G182269; -.  
DR PIR: A32066; TVHUET.  
DR PIR: A10086; S10086.  
DR PDB: 2STT; 12-MAR-97.  
DR PDB: 2STW; 12-MAR-97.  
DR TRANSFAC: T00112; -.  
DR MIM: 164720; -.  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
KW PROTO-ONCOGENE; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING;  
KW PHOSPHORYLATION; 3D-STRUCTURE.  
ET DNA\_BIND 335 415 ETS-DOMAIN.  
ET VARSPIC 244 330 MISSING (IN C-ETS-1B).  
SQ SEQUENCE 441 AA; 50408 MM; 958AF4A4 CRC32;  
Query Match 14.9%; Score 490; DB 1; Length 441;  
Best Local Similarity 59.0%; Pred. No. 2,33e-83;  
Matches 62; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

